One

OVERVIEW

Despite the plethora of IQ tests available for psychologists to use today, the Wechsler instruments remain the most widely used measures of intelligence for children, adolescents, and adults. Much has been written on these measures over the years, from clinical use of the scales to esoteric statistical procedures for interpreting the profiles that they yield. Our goal for this book is to provide an easy reference source for those who use the Wechsler Preschool and Primary Scale of Intelligence–Third Edition (WPPSI-III; The Psychological Corporation, 2002). This book was developed for those who test children within the 2-1/2 to 7-year age range and wish to learn the essentials of the WPPSI-III in a direct, no-nonsense, systematic manner. The main topics covered include administration, scoring, interpretation, and clinical use of the instrument. Important points are highlighted throughout the book by Rapid Reference boxes, Caution boxes, and Don't Forget boxes. Each chapter contains questions that are intended to help you consolidate what you have read. After reading this book, you will have, at your fingertips, in-depth information that will help you to become a competent WPPSI-III examiner and clinician.

HISTORY AND DEVELOPMENT

Although interest in testing intelligence developed in the latter half of the 19th century, the assessment of preschool-age children is a relative newcomer in the history of testing (Kelley & Surbeck, 2000). In the early 1900s, the majority of tests were developed for school-age children, leaving a hole in the area of preschool measures.

Shortly after the end of the 19th century, Alfred Binet and his colleagues developed tasks to measure the intelligence of children within the Paris

public schools (Binet & Simon, 1905). Binet's tasks were primarily language oriented, emphasizing judgment, memory, comprehension, and reasoning. In the 1908 revision of his scale, Binet included age levels ranging from 3 to 13 years; and in its next revision in 1911, the Binet-Simon scale was extended to age 15 and included five ungraded adult tests (Kaufman, 1990a). Kuhlmann (1912, 1914) published two versions of the Binet scales, the second of which extended test items downward to assess intelligence beginning at 2 months of age. Although the versions of intelligence tests published by Kuhlmann (1914), Yerkes and Foster (1923), and Burt (1921) increased attention to assessment of preschoolers, these early tests were methodologically lacking (Stott & Ball, 1965).

Gesell (1925) subsequently undertook a seminal study in child development. Children were examined at 10 age levels — birth, 4, 6, 9, 12, 18, 24, 36, 48, and 60 months. Although precise methodology was not used, the study yielded "developmental schedules" across four areas: motor development, language development, adaptive behavior, and personal-social behavior. The developmental profiles derived from Gesell's work were subsequently used in the development of tests for infants and preschoolers.

Key assessment instruments for measurement of infant and preschool development were published in the first half of the 20th century. Most notable were the Merrill-Palmer Scale of Mental Tests (Stutsman, 1931), the Minnesota Preschool Scale (Goodenough, 1926; Goodenough, Maurer, & Van Wagenen, 1940), the California First Year Mental scale (Bayley, 1933), and the Iowa Test for Young Children (Fillmore, 1936). These early infant and preschool tests focused more on mental and physical growth than on intelligence.

The 1940s saw many new tests published for infant and preschool assessment, most notably the Cattell Infant Intelligence Scale (Cattell, 1940), the Northwest Infant Intelligence Scale (Gilliland, 1948), the Leiter International Performance Scale (Leiter, 1948), and the Full Range Picture Vocabulary Test (Ammons & Ammond, 1948). Although these tests made unique contributions to the field of preschool assessment (e.g., the Leiter was a nonlanguage, allegedly culture-free test and the Full Range Picture Vocabulary tests had high reliability and validity), the Stanford-Binet continued to be the most widely used test of mental ability (Goodenough, 1949).

The Stanford-Binet, however, had some major competition after David Wechsler's tests entered the playing field in the mid-1930s. Wechsler's

approach combined his strong clinical skills and statistical training with his extensive experience in testing, gained initially as a World War I examiner. Wechsler weighted verbal and nonverbal abilities equally, an innovative idea at that time. Wechsler's goal was to create a battery that would yield dynamic clinical information from his chosen set of tasks. This focus went well beyond the earlier use of tests simply as psychometric tools. Wechsler's first test for children, the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1949), was a downward extension of Form II of the Wechsler Bellevue (Wechsler, 1946) and covered the age range of 5–15 years. Years later, the WISC became one of the most frequently used tests in the measurement of preschool functioning (Stott & Ball, 1965), although it was not able to be used with children below age 5. The practice of using tests designed for school-aged children in assessing preschoolers was criticized because of the level of difficulty for young children; nonetheless, the downward extension of tests designed for school-aged children was common practice prior to the development of tests specifically geared for children under age 5 (Kelley & Surbeck, 2000).

The primary focus of the testing movement prior to the 1960s was the assessment of children in school and of adults entering the military (Parker, 1981). However, in the 1960s, the U.S. federal government began to play a role in education, and this involvement spurred growth in the testing of preschool children. The development of government programs such as Head Start focused attention on the need for effective program evaluation and the adequacy of preschool assessment instruments (Kelley & Surbeck, 1991). In 1967 the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) was developed to meet the growing need of how to evaluate programs such as Head Start. The WPPSI was basically developed as a downward extension of many of the WISC subtests, but it provided simpler items and an appropriately aged standardization sample. Unfortunately, the WPPSI accommodated the narrow 4- to 6 1/2-year age range, failing to meet the needs of program evaluations because most of the new programs were for ages 3 to 5 years.

Shortly after the WPPSI, the McCarthy Scales of Children's Abilities (MSCA; McCarthy, 1972) was published. The McCarthy was based on normative data gathered on 1,032 children ages 2 1/2 through 8 1/2 years. The unique features of the McCarthy made it valuable for the assessment of

children with learning problems or other exceptionalities. The McCarthy yielded not only a general measure of intellectual functioning but also a profile of abilities including verbal ability, nonverbal reasoning, number aptitude, short-term memory, and motor coordination.

Public Law 94-142, the Education for All Handicapped Children Act of 1975, played an important role in the continued development of cognitive assessment instruments. This law and its followers (Individuals with Disabilities Education Act [IDEA], IDEA of 1991, and IDEA Amendments in 1997) included provisions requiring that an individualized education program (IEP) be developed and maintained for each disabled child (Kelley & Surbeck, 2000). A key feature of the development of the IEP is the evaluation and diagnosis of each child's level of functioning. Thus, these laws directly affected the continued development of standardized tests such as the WPPSI. The WPPSI has had two revisions — one in 1989 and its most recent in 2002. The Don't Forget box on page 5 shows the history of Wechsler's scales.

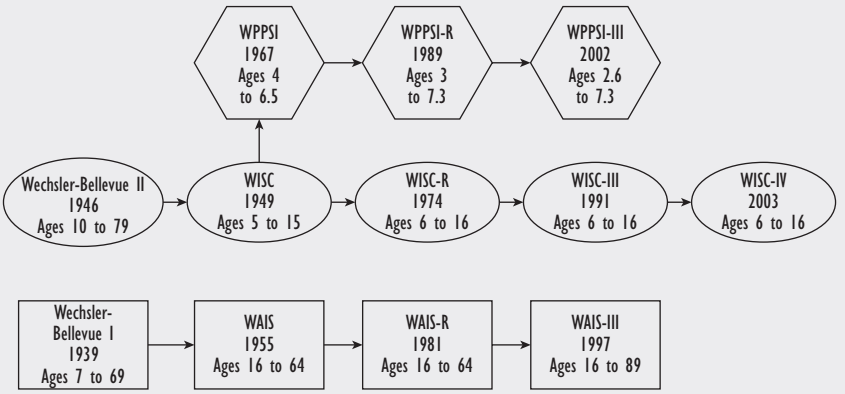
THEORETICAL FOUNDATION

Historically, the concept of intelligence has been difficult to define, and even today it remains elusive (Flanagan, Genshaft, & Harrison, 1997). Wechsler's (1944) conception of intelligence as "the capacity to act purposefully, to think rationally, and to deal effectively with his [or her] environment" (p. 3) provided the foundation of all Wechsler tests, including the current editions. Practical and clinical perspectives were the cornerstone of Wechsler's tests rather than theory per se (except, perhaps, for Spearman's *g* or general intelligence theory). However, test developers at The Psychological Corporation created some of the newest WPPSI-III subtests to update the test's theoretical foundations. The origin of each of the WPPSI-III subtests is shown in Rapid Reference 1.1.

Like the WISC-III and WAIS-III, the third edition of the WPPSI contains subtests that were designed to tap more specific theoretically-based abilities, such as processing speed and fluid reasoning. Fluid reasoning is a specific cognitive ability that has been emphasized by several theorists (Carroll, 1997; Cattell, 1941, 1963; Cattell & Horn, 1978; Horn & Noll, 1997). Fluid reasoning tasks involve the process of "manipulating abstractions, rules, generalization, and logical relationships" (Carroll, 1993, p. 583). Three new subtests

DON'T FORGET

History of Wechsler Intelligence Scales



Note. WISC = Wechsler Intelligence Scale for Children; WAIS = Wechsler Adult Intelligence Scale.

were added to the WPPSI-III to enhance the measurement of fluid reasoning: Matrix Reasoning, Word Reasoning, and Picture Concepts. Carroll (1993) and other theorists (e.g., Horn & Noll, 1997) also identified processing speed as an important domain of cognitive functioning. Thus, two new subtests measuring processing speed were added to the WPPSI-III battery, namely Symbol Search and Coding.

Although the newest version of the WPPSI has increased its emphasis on the importance of theoretical foundations, originally Wechsler believed that IQ tests offered a way to peer into an individual's personality. Since the development of the Wechsler scales, extensive theoretical speculations have been made about the nature and meaning of these tests and their scores (Kaufman, 1990a, 1994b), but originally the tests were developed without regard to theory. The Wechsler tests are strongly supported as measures of general intelligence (*g;* e.g., Kaufman, 1994b), but — as we show throughout this book — much more can be gleaned from the Wechsler scales than simply an understanding of a child's level of *g*.

Wechsler made a major contribution to the fields of cognitive and clinical assessment with his inclusion of both Verbal and Performance Scales on his

Rapid Reference 1.1

Origin of WPPSI-III Subtests

Verbal Subtest	Historical Source of Subtest
Vocabulary	Stanford-Binet
Similarities	Stanford-Binet
Information	Army Alpha
Comprehension	Stanford-Binet, Army Alpha
Word Reasoning	Kaplan's Word Context Test (Werner & Kaplan, 1950)
Receptive Vocabulary	Stanford-Binet
Picture Naming	Stanford-Binet
Performance Subtest	**Historical Source of Subtest**
Picture Completion	Army Beta, Army Performance Scale Examination
Coding	Army Beta, Army Performance Scale Examination
Block Design	Kohs (1923)
Matrix Reasoning	Raven's Progressive Matrices (1938)
Symbol Search	Shiffrin & Schneider (1977) and S. Sternberg (1966)
Object Assembly	Army Performance Scale Examination
Picture Concepts	Novel task developed by Psychological Corporation

tests. The dual-scaled tests went against the conventional wisdom of his time. In the 1930s and 1940s, it didn't make sense to most examiners to waste their time administering a lengthy nonverbal subtest when a quick verbal subtest could glean just as much data. However, now it is obvious to clinicians and researchers alike that Verbal and Performance both have critical value for understanding brain functioning and theoretical distinctions between fluid and crystallized intelligence. In addition, because Wechsler stressed the clinical value of intelligence tests, this innovative approach provided a new layer

to the psychometric, statistical emphasis of testing that accompanied the use and interpretation of earlier tests such as the Stanford-Binet. Finally, Wechsler's inclusion of a multiscore subtest profile (as well as three IQs instead of one) met the needs of the emerging field of learning disabilities assessment in the 1960s to such an extent that Wechsler's scales replaced the Stanford-Binet as king of IQ during that decade. It has maintained that niche ever since.

PURPOSES OF ASSESSING PRESCHOOLERS AND SCHOOL-AGE CHILDREN

Children are assessed for a variety of reasons; thus, the WPPSI-III may be applied in many different situations. Typically, children are referred by a teacher for a psychological evaluation to determine whether they are eligible for an educationally related disability and special education or other special services. Some of the most common reasons that a child is referred for an assessment include diagnosing for developmental delay, learning disabilities, mental retardation, behavioral problems, neuropsychological impairments, or giftedness. Often, the end goal of a child's assessment is to create effective interventions. The number of children ages 3 to 5 years in the United States who were served in federally supported programs for persons with disabilities (including specific learning disabilities, mental retardation, developmental delay, and other disabilities) numbered nearly 600,000 in 1999–2000 (U.S. Department of Education, 2001). The settings in which these assessments take place are varied and include psychologists' private practices, schools, clinics, hospitals, and research programs.

As mentioned earlier, the Wechsler scales remain by far the most popular test for children (Daniel, 1997). In a survey of school psychologists who assess children to identify mental retardation, the Wechsler scales were the most frequently used tests for deriving IQs (Woodrich & Barry, 1991). Even in assessing children with bilingual and limited-English students, the WISC-R and WISC-III were reported to be the most frequently used measures (Ochoa, Powell, & Robles-Pina, 1996). School psychologists rated the Wechsler scales as most useful and as actually used the most (Giordano, Schwiebert, & Brotherton, 1997), and in another survey of school psychologists, the WISC-III was reportedly used 10 times per month, whereas the next most frequently used test (of 11 listed) was used only twice (Wilson & Reschly,

1996). Because of the Wechsler scales' popularity throughout the years, the WPPSI and WPPSI-R have remained strong forces in the assessment of preschool-aged children, and the WPPSI-III is sure to follow suit.

DESCRIPTION OF WPPSI-III

The WPPSI-III is a measure of cognitive functioning of children from ages 2 years, 6 months (2-6) to 7 years, 3 months (7-3). Its age range is divided into two age bands (2-6 to 3-11 and 4-0 to 7-3), each with its own battery of subtests. Like its predecessors, the WPPSI-III offers a Verbal IQ (V-IQ), Performance IQ (P-IQ), and Full Scale IQ (FS-IQ). However, departing from the previous versions of the WPPSI, the WPPSI-III adds a General Language Composite (GLC) and — for the older age band — a Processing Speed Quotient (PSQ) to the three familiar IQs. Like the IQs, the GLC and PSQ are standard scores with a mean of 100 and standard deviation of 15. Mainly motor responses are required on the Performance scale (pointing, placing, or drawing), and spoken responses are usually required on the Verbal scale.

For each age band, WPPSI-III subtests are categorized as core, supplemental, or optional. Core subtests are those that comprise the V-IQ, P-IQ, and FS-IQ. The composition of the scales for each age group is presented in Figures 1.1 and 1.2. In the younger age bracket, two core subtests comprise the V-IQ and two comprise the P-IQ. The four subtests of the V-IQ and P-IQ together yield the FS-IQ for children ages 2–6 to 3–11. In the older age bracket, three subtests comprise the V-IQ and three comprise the P-IQ. In addition to the six subtests of the V-IQ and P-IQ, an additional core subtest (Coding) is added in the calculation of the FS-IQ for those aged 4–0 to 7–3. For both age groups, the GLC comprises two subtests: Receptive Vocabulary and Picture Naming (a supplemental subtest for children under age 4 and an optional one for those age 4 and above). Only the older age bracket has a fourth standard score, the PSQ, which is composed of Coding and Symbol Search (a supplemental subtest). Because GLC and PSQ require the administration of noncore subtests, these two global scores are supplements, not core standard scores. Consistent with the metric used for all Wechsler subtests, each WPPSI-III subtest yields a scaled score with a mean of 10 and a standard deviation of 3. Rapid Reference 1.2 lists and describes each WPPSI-III subtest.

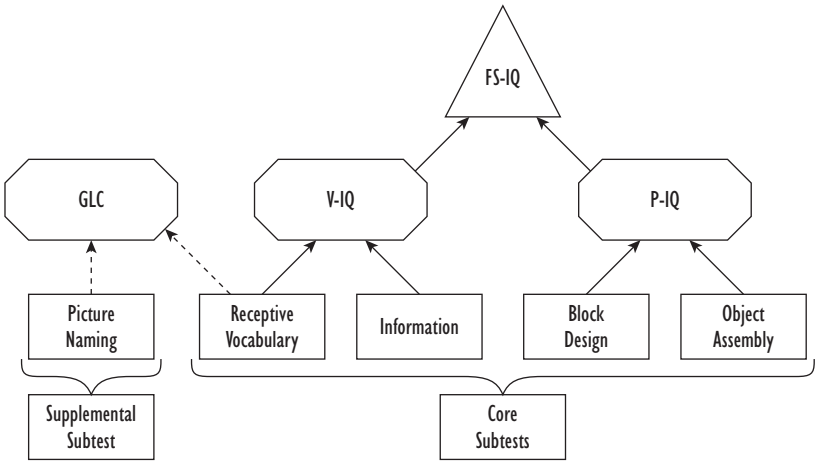


Figure 1.1 WPPSI-III building blocks for Ages 2-6 to 3-11.

Note. FS-IQ = Full Scale IQ; V-IQ = Verbal IQ; P-IQ = Performance IQ; GLC = General Language Composite. Picture Naming can be substituted for a core Verbal subtest if necessary.

CHANGES FROM WPPSI-R TO WPPSI-III

With their revision of the WPPSI-R, the professionals at The Psychological Corporation intended to improve the psychometric properties, strengthen the test's theoretical foundations, enhance its clinical utility, increase the age appropriateness, and enhance the user friendliness of the test. We believe that they achieved their goals. Rapid Reference 1.3 lists the five WPPSI-III revision goals and how those goals were met.

Significant changes in the composition of the scales were made when the WPPSI-R was transformed into the WPPSI-III. Most notably, five WPPSI-R subtests were dropped (Arithmetic, Animal Pegs, Geometric Design, Mazes, and Sentences), and seven new subtests were added: Receptive Vocabulary, Picture Naming, Word Reasoning, Matrix Reasoning, Picture Concepts, Coding, and Symbol Search. The Psychological Corporation (2002) stated that the five deleted subtests were removed in part because the total number of subtests was too great with the additional seven new tasks. The deleted subtests were all influenced by factors other than intellectual capability, including neurological and motor development, as well as familiarity with numbers and abstract concepts. The eliminated subtests were also those that tapped memory capabilities of young children. The creators of the WPPSI-III recognize

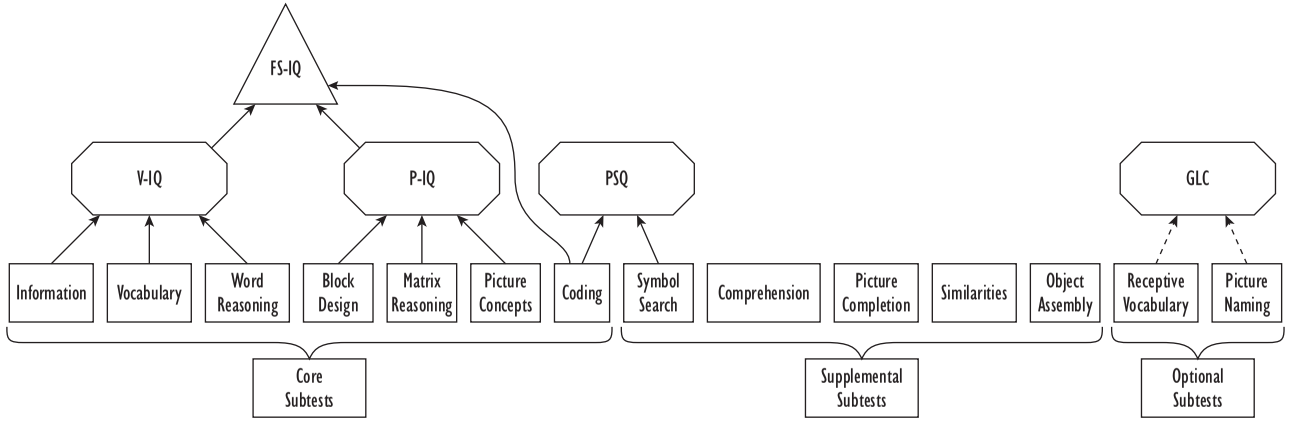


Figure 1.2 WPPSI-III building blocks for Ages 4-0 to 7-3.

Note. FS-IQ = Full Scale IQ; V-IQ = Verbal IQ; P-IQ = Performance IQ; PSQ = Processing Speed Quotient; GLC = General Language Composite. Either supplemental Verbal subtest may be substituted for one core Verbal subtest. Supplemental Performance subtests may be substituted for one core Performance subtest.

Rapid Reference 1.2

Description of WPPSI-III Subtests

Performance Subtests

Retained from WPPSI-R

Object Assembly. Child is required to fit puzzle pieces together to form a meaningful whole.

Block Design. Child reproduces patterns made from 1- or 2-colored blocks.

Picture Completion. Child identifies what is missing from pictures of common objects.

Newly Developed for WPPSI-III

Matrix Reasoning. The child looks at an incomplete matrix and selects the missing section from four or five response options.

Picture Concepts. Child is presented with two or three rows of pictures and chooses one picture from each row to form a group with a common organizational concept.

Symbol Search. Child indicates, by marking a box, whether a target symbol appears in a series of symbols.

Coding. Using a key, the child draws symbols that are paired with simple geometric shapes.

Verbal Subtests

Retained from WPPSI-R

Information. Child must either point to a picture or verbally answer brief oral questions about commonplace objects and events.

Comprehension. Child verbally responds to questions about consequences of events.

Vocabulary. Child names pictured items and provides verbal definitions of words.

Similarities. Child completes a sentence that contains a verbal analogy.

Newly Developed for WPPSI-III

Picture Naming. The child names pictures that are displayed singularly in the stimulus booklet.

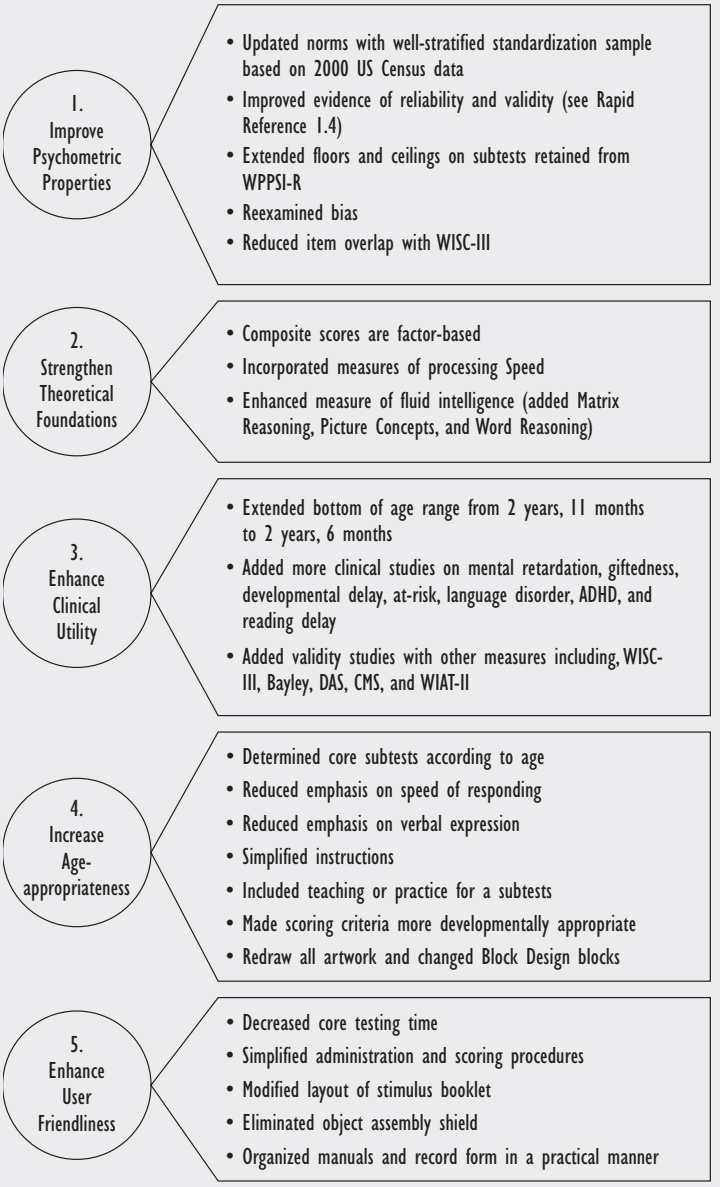
Receptive Vocabulary. The child looks at a group of four pictures and points to the one that the examiner describes aloud.

Word Reasoning. The child is read an increasingly specific series of one to three clues and identifies the common object or concept being described.

Note. Subtests that were retained from the WPPSI-R have new items and contain changes in administration and scoring.

Rapid Reference 1.3

Revision Goals of the WPPSI-III



the importance of memory for young children (The Psychological Corporation, 2002, p. 23) and suggest that examiners administer a comprehensive test of memory if it is warranted.

With the extra subtests in the WPPSI-III came the two additional composite scores, the GLC and PSQ. These new composites follow well from the factorial and theoretical structure of the scale. Rapid Reference 1.4 shows the relationships between old WPPSI-R and current WPPSI-III scales and subtests. Correlations were strongest between the WPPSI-R and WPPSI-III Information (.80) and Vocabulary (.77) subtests, each of which endured only minor changes. Four items were added to Vocabulary and 16 were retained, and on Information, six new picture items were added along with nine verbal items (19 verbal items were retained). The weakest relationships were for Object Assembly (.53) and Similarities (.57), both of which were modified substantially. Object Assembly retained only two WPPSI-R items and added 12 new ones; administration and scoring procedures were also changed.

Rapid Reference 1.4

Correlations Between WPPSI-III and the WPPSI-R Subtests of the Same Name and IQ Scales of the Same Name

Subtest or IQ Scale	Corrected *r*
Information	.80
Vocabulary	.77
Comprehension	.68
Picture Completion	.66
Block Design	.57
Similarities	.57
Object Assembly	.53
Verbal IQ	.86
Performance IQ	.70
Full Scale IQ	.85

Note. All values are corrected using Fisher's *z* transformation. *N*s varied across individual subtests and ranged from 129 to 176. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Table 5.7).

Similarities was modified by removing two of the three WPPSI-R item types and adding 16 new items (eight old items were retained).

STANDARDIZATION AND PSYCHOMETRIC PROPERTIES OF WPPSI-III

Standardization

The WPPSI-III was standardized on a sample of 1,700 children who were chosen to match closely the 2000 U.S. Census data on the variables of age, gender, geographic region, ethnicity, and parental education. The standardization sample was divided into nine age groups, each composed of 200 children, except for the 7-0 to 7-3 age group that was composed of 100 children. The sample was split equally between boys and girls.

Reliability

The reliability and validity information is presented in the *WPPSI-III Technical and Interpretive Manual* (The Psychological Corporation, 2002) and is summarized in Rapid Reference 1.5. The average internal consistency coefficients are .95 for V-IQ, .93 for P-IQ, .89 for PSQ, .93 for GLC, and .96 for FS-IQ. Internal consistency values for individual subtests across all ages ranged from 0.75 for Block Design (for the 4-0 to 4-5 age group) to .96 on Similarities (for ages 4-6 to 4-11 and 5-6 to 5-11). The median internal consistency value for the individual subtests was 0.88.

The WPPSI-III is a fairly stable instrument with average test-retest coefficients of 0.91, 0.86, and 0.92 for the V-, P-, and FS-IQ, respectively (see Rapid Reference 1.5 for a reliability summary that includes internal consistency and stability values). The stability values of the PSQ (.86) and GLC (.91) were consistent with the coefficients for the IQs with the P-IQ and PSQ emerging as the least stable of all the composite scores. The largest practice effects (i.e., score increases from first testing to second) for the combined age bands were 5–6 points for P-IQ and PSQ. The average practice effects for all ages for V-IQ and GLC were just under 3 points (see Rapid Reference 1.6).

Rapid Reference 1.5

Average WPPSI-R and WPPSI-III Reliability Coefficients

WPPSI-R/WPPSI-III Scale or Subtest	WPPSI-R Split-half reliability	WPPSI-R Test-retest reliability	WPPSI-III Split-half reliability	WPPSI-III Test-retest reliability
Verbal IQ	.95	.90	.95	.91
Performance IQ	.92	.88	.93	.86
Full Scale IQ	.96	.91	.96	.92
Processing Speed Quotient	—	—	.89	.90
General Language Composite	—	—	.93	.91
Verbal				
Information	.84	.81	.88	.86
Similarities	.86	.70	.95	.90
Arithmetic	.80	.71	—	—
Vocabulary	.84	.75	.89	.84
Comprehension	.83	.78	.88	.81
Picture Naming	—	—	.88	.88
Receptive Vocabulary	—	—	.88	.83
Word Reasoning	—	—	.91	.82
Sentences	.82	.79	—	—
Performance				
Picture Completion	.85	.82	.90	.85
Coding	—	—	—	.84
Matrix Reasoning	—	—	.90	.81
Block Design	.85	.80	.84	.76
Object Assembly	.63	.59	.85	.74
Symbol Search	—	—	—	.83
Picture Concepts	—	—	.91	.75
Mazes	.77	.52	—	—
Geometric Design	.79	.67	—	—
Animal Pegs[a]	—	.66	—	—

[a]For Animal Pegs, Coding, and Symbol Search, and for the Processing Speed Quotient, only test-retest coefficients are reported because of the timed nature of the subtests.

The youngest age band (2-6 to 3-11) had smaller practice effects than the older two bands (4-0 to 5-5 and 5-6 to 7-3). Like P-IQ, the individual Performance subtests generally had larger gains in scaled scores on retesting than did the Verbal subtests. Rapid Reference 1.7 shows the subtests that have relatively large gains from test to retest (large gains are defined as those at least 0.9 scaled score points, which equals 0.3 of a standard deviation).

Validity

Construct validity of the WPPSI-III is supported by the factor-analytic studies described in the *Technical and Interpretive Manual.* For the 2-6 to 3-11 age group, the WPPSI-III is a two-factor test, Verbal and Performance. For ages 4-0 to 7-3, a third factor emerges, Processing Speed. When only the core subtests were analyzed, the WPPSI-III subtests each loaded on its predicted factor, with the exception of Picture Concepts. At every age level except 6-0 to 7-3, Picture Concepts loaded on its intended factor — Performance — but in the oldest age range, Picture Concepts was more decisively a Verbal than Performance subtest.

Rapid Reference 1.6

Practice Effects for the WPPSI-III Global Scales

Scale	Ages 2.6-3.11 (*N* = 41)	Ages 4.0-5.5 (*N* = 34)	Ages 5.6-7.3 (*N* = 82)	All Ages (*N* = 157)
V-IQ	+1.6	+5.0	+2.4	+2.8
P-IQ	+4.2	+4.8	+5.7	+5.0
FS-IQ	+3.4	+6.4	+5.5	+5.2
GLC	+3.0	+5.2	+1.5	+2.7
PSQ	—	+6.0	+6.3	+6.2

Note. Data are from *WPPSI-III Technical and Interpretive Manual* (Table 4.4). Intervals ranged from 14 to 50 days with a mean of 26 days.

Rapid Reference 1.7

Practice Effects for the Separate WPPSI-III Scaled Scores: Subtests With Relatively Large Gains From Test to Retest

Ages 2.6-3.11 (*N* = 41)	Ages 4.0-5.5 (*N* = 34)	Ages 5.6-7.3 (*N* = 82)
Block Design (+0.9)	Picture Completion (+1.5)	Picture Completion (+1.4)
	Matrix Reasoning (+1.1)	Matrix Reasoning (+1.3)
	Coding (+1.1)	Coding (+1.1)
	Symbol Search (+1.1)	Symbol Search (+1.0)
	Object Assembly (+1.1)	Similarities (+0.9)
	Word Reasoning (+1.0)	
	Receptive Vocab. (+1.0)	
	Information (+0.9)	
	Picture Naming (+0.9)	

Note. Relatively large gains are defined as at least 0.3 *SD* (a gain of at least 0.9 scaled-score points from test to retest). Data are from *WPPSI-III Technical and Interpretive Manual* (Table 4.4). Intervals ranged from 14 to 50 days with a mean of 26 days.

When factor analyses included WPPSI-III supplemental as well as core subtests, the results were not as consistent with predictions. Picture Concepts was, again, a maverick subtest. As shown in Table 1.1, it loaded equally on both Processing Speed and Verbal at ages 4-0 to 4-11, it loaded primarily on Performance at ages 5-0 to 5-11, and it was primarily a Verbal subtest at ages 6-0 to 7-3.

Table 1.1 Factor Loadings for Picture Concepts Derived from Exploratory Analysis With Core and Supplemental WPPSI-III Subtests

Age	Verbal	Performance	Processing Speed
4-0 to 4-11	*.33*	.07	*.34*
5-0 to 5-11	.15	*.46*	.15
6-0 to 7-3	*.51*	.10	.06

Note. Loadings > .30 are in italics. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Table 5.4).

Table 1.2 Subtests With Four Highest and Four Lowest Correlations with Picture Concepts

Subtest	*r*
Highest	
Word Reasoning	.51
Similarities	.51
Information	.49
Matrix Reasoning	.48
Lowest	
Picture Naming	.42
Block Design	.41
Object Assembly	.39
Coding	.32

Note. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Table 5.1).

Table 1.3 Factor Loadings for Matrix Reasoning Derived From Exploratory Analysis With Core and Supplemental WPPSI-III Subtests

Age	Verbal	Performance	Processing Speed
4-0 to 4-11	.16	.19	*.39*
5-0 to 5-11	.12	*.36*	*.30*
6-0 to 7-3	.26	*.59*	-.09

Note. Loadings > .30 are in italics. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Table 5.4).

Like Picture Concepts, Matrix Reasoning also loaded on multiple factors in the exploratory factor analyses (see Table 1.3). At ages 4-0 to 4-11, Matrix Reasoning loaded only on Processing Speed, and it loaded about equally on Processing Speed and Performance at ages 5-0 to 5-11. At 6-0 to 7-3, it loaded only on the Performance factor. Therefore, the age trends suggest that

Table 1.4 Matrix Reasoning's Correlations With Processing Speed Subtests Across Six Separate Age Groups

	Correlations With Matrix Reasoning						
	Ages 4-0 to 4-5	Ages 4-6 to 4-11	Ages 5-0 to 5-5	Ages 5-6 to 5-11	Ages 6-0 to 6-11	Ages 7-0 to 7-3	Median
Symbol Search	.44	.52	.52	.48	.50	.50	.50
Coding	.41	.44	.55	.40	.22	.12	.41

Note. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Tables A.4, A.5, A.6, A.7, A.8, and A.9).

as children progress from age 4 to 7 years, Matrix Reasoning becomes increasingly a function of Performance ability. Possible explanations of the age trends for Picture Concepts and Matrix Reasoning appear in chapter 4.

In addition to factor analyses, validity of the WPPSI-III is further supported by correlations with the following instruments (The Psychological Corporation, 2002): Bayley Scales of Infant Development–II (BSID-II; Bayley, 1993), WPPSI-R (Wechsler, 1989), WISC-III (Wechsler, 1991), and Differential Abilities Scale (DAS; Elliott, 1990). Each of the global scales of these four instruments correlated strongly with the WPPSI-III FS-IQ. Correlations ranged from .80 to .89 (see Rapid Reference 1.8). Rapid Reference 1.9 also shows that the WPPSI-III Verbal Scale correlated substantially higher with the verbal scales of the WPPSI-R, WISC-III, and DAS than it did with the nonverbal scales of each instrument. These patterns of correlations support the convergent and discriminant validity of the WPPSI-III. Chapter 5 presents a more detailed review of validity issues, and chapter 6 touches on the validity of the WPPSI-III in special populations.

To evaluate the relationship of the WPPSI-III scores to the key criterion of academic achievement (The Psychological Corporation, 2002), 208 children were administered both the WPPSI-III and the Wechsler Individual Achievement Test–Second Edition (WIAT-II; The Psychological Corporation, 2001).

Rapid Reference 1.8

Correlations of WPPSI-III Full Scale IQ With Other Global Measures

	WPPSI-III FS-IQ
Bayley Scales of Infant Development–II (BSID-II) (*N* = 84)	
Mental Score	.80
WPPSI-R (*N* = 176)	
Full Scale IQ	.85
WISC-III (*N* = 94)	
Full Scale IQ	.89
Differential Ability Scales (DAS) (*N* = 153)	
General Conceptual Ability (GCA) Standard Score	.87

Note. All values are corrected for the variability of the standardization sample. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Tables 5.7, 5.9, 5.11, and 5.13).

The strongest correlation was between the WPPSI-III FS-IQ and WIAT-II Total Achievement (.78) and the weakest was between the PSQ and Reading (.31). The coefficients between WPPSI-III and WIAT-II global scores are presented in Rapid Reference 1.10. As shown, V-IQ correlated strongly with Total Achievement (.77), and the P-IQ correlated substantially with the Mathematics Composite (.60). The validity of new WPPSI-III Verbal subtests was also supported with correlations to WIAT-II composites (see Rapid Reference 1.11). Picture Naming (.71) and Word Reasoning (.70) were among the best correlates of Total Achievement. New Performance subtests were not as strongly correlated with Achievement: Matrix Reasoning and Picture Concepts both correlated .35 with Total Achievement. However, as shown in Rapid Reference 1.11, the old subtests tended to be both the best and worst

correlates of achievement on the WIAT-II. Information and Similarities were the highest correlates of Mathematics, Oral Language, and Total Achievement. Picture Completion, Object Assembly, and Coding were the lowest correlates of Reading and Total Achievement. Overall, the WPPSI-III–WIAT-II relationships replicate prior research and support the validity of the WPPSI-III.

Rapid Reference 1.9

Convergent and Discriminant Validity of the WPPSI-III Verbal-Performance IQ Discrepancy: Correlations of WPPSI-III V-IQ and P-IQ with Other Measures of Verbal and Nonverbal Ability

	WPPSI-III V-IQ	WPPSI-III P-IQ
WPPSI-R (*N* = 176)		
V-IQ	**.86**	.59
P-IQ	.60	**.70**
WISC-III (*N* = 96)		
V-IQ	**.82**	.67
P-IQ	.60	**.79**
DAS (*N* = 112)		
Verbal	**.78**	.54
Nonverbal Reasoning	.56	**.76**

Note. Coefficients in bold denote convergent validity of WPPSI-III Verbal and Performance IQs. All values are corrected for the variability of the standardization sample. Most values are from *WPPSI-III Technical and Interpretive Manual* (Tables 5.7, 5.9, and 5.13). The remaining values were kindly provided by J. J. Zhu (personal communication, October 23, 2002).

Rapid Reference 1.10

WPPSI-III IQs and the Processing Speed Index: Correlations With WIAT-II Achievement Composites

WIAT-II Composite	*N*	Verbal IQ	Perform-ance IQ	Full Scale IQ	Processing Speed Index
Reading	58	.60	.44	.66	.31
Math	133	.56	.60	.77	.55
Written Language	58	.59	.36	.62	.41
Oral Language	201	.72	.44	.67	.39
Total	56	.77	.55	.78	.36

Note. All values are corrected for the variability of the standardization sample. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Table 5.14).

COMPREHENSIVE REFERENCES ON THE TEST

The *WPPSI-III Technical and Interpretive Manual* (The Psychological Corporation, 2002) provides detailed information about the development of the test; descriptions of the subtests and scales; and information about the test's standardization, reliability, and validity. We have found no other comprehensive references on the WPPSI-III. There are, however, several comprehensive treatments of its predecessor, the WPPSI-R. Chapter 11 of Sattler's (2001) *Assessment of Children: Cognitive Applications (4th edition)* presents an overview of what the test measures and an approach to interpretation. Gyurke's (1991) chapter on the WPPSI-R describes the subtests and scales, summarizes psychometric information, and provides steps for interpreting the test. Kaufman and Lichtenberger's (2000) *Essentials of WISC-III and WPPSI-R Assessment* provides the same type of treatment of WPPSI-R administration, scoring, interpretation, and applications that is detailed in the present book for the WPPSI-III. Rapid Reference 1.12 provides basic information on the WPPSI-III and its publisher.

Rapid Reference 1.11

WPPSI-III Subtests: The Highest and Lowest Correlates of WIAT-II Achievement Composites

Correlations of WPPSI-III Scaled Scores with WIAT-II Achievement Composite Standard Scores

Reading (*N* = 58)	Written Math (*N* = 133)	Oral Language (*N* = 58)	Language (*N* = 201)	Total (*N* = 56)
Highest	**Highest**	**Highest**	**Highest**	**Highest**
I (60)	I (54)	WR (57)	I (67)	I (72)
PN (60)	S (54)	C (57)	S (63)	WR (71)
RV (59)	RV (51)	RV (54)	WR (62)	S (70)
	BD (51)			PN (70)
Lowest	**Lowest**	**Lowest**	**Lowest**	**Lowest**
Cd (24)	OA (37)	PC (31)	OA (26)	PC (32)
PC (29)	MR (36)	MR (22)	MR (24)	OA (26)
OA (16)	WR (35)	OA (16)	Cd (19)	Cd (25)

Note. Decimal points are omitted. All values are corrected for the variability of the standardization sample. Within each column, coefficients are listed from high to low. So, for example, the best predictor of WIAT-II Reading is Information and the worst is Object Assembly. Coefficients are from *WPPSI-III Technical and Interpretive Manual* (Table 5.14).

Note. I = Information; V = Vocabulary; WR = Word Reasoning; C = Comprehension; S = Similarities; RV = Receptive Vocabulary; PN = Picture Naming; BD = Block Design; MR = Matrix Reasoning; PCon = Picture Concepts; SS = Symbol Search; Cd = Coding; PC = Picture Completion; OA = Object Assembly.

Rapid Reference 1.12

Basic Information About the Wechsler Preschool and Primary Scale of Intelligence–Third Edition

Author	The Psychological Corporation
Publication date	2002
What the test measures	Verbal, nonverbal, and general intelligence, processing speed, and general language abilities
Age range	2 years, 6 months to 7 years, 3 months
Administration time	Ages 2 years, 6 months to 3 years, 11 months: 30–45 min Ages 4 years to 7 years, 3 months: 60 min
Qualification of examiners	Graduate- or professional-level training in psychological assessment
Publisher	The Psychological Corporation 555 Academic Court San Antonio, TX 78204-2498 800-211-8378 http://www.PsychCorp.com
Price (from 2003 catalog)	**WPPSI™–III Kit** Includes all necessary stimulus and manipulative materials, Examiner Manual, Technical Manual, 25 Record Forms for ages 2-6 to 7-3, 25 Record Forms for ages 2-6 to 3-11, and 25 Response Booklets $725.00 (in box) or $775 (in attaché or soft-sided case) **WPPSI™–III Scoring Assistant®** CD-ROM Windows® $165.00 **WPPSI™–III Writer™** CD-ROM Windows® $350.00

TEST YOURSELF

1. **Prior to the development of the Wechsler scales, the Stanford-Binet was the most widely used test of mental ability for preschoolers.** True or False?
2. **Prior to the development of the WPPSI, which Wechsler test was commonly administered to preschoolers?**
   (a) WAIS
   (b) Wechsler Bellevue-II
   (c) WISC
   (d) WIAT
3. **What law played an important role in the continued development of cognitive assessment instruments for children?**
4. **Preschool assessment measures such as the WPPSI-III were developed as upward extensions of infant tests of intelligence.** True or False?
5. **The WPPSI-III P-IQ significantly correlated with which achievement measure?**
   (a) WIAT-II Listening Comprehension
   (b) WIAT-II Oral Language
   (c) WIAT-II Mathematics
   (d) WIAT-II Reading
6. **Given the results of the factor analyses of the WPPSI-III, you should not be surprised if a 6-year-old's score on Picture Concepts (a Performance subtest) is more similar to scores on Verbal subtests than to scores on other Performance subtests.** True or False?
7. **The two new composites added to the WPPSI-III are**
   (a) Processing Speed Quotient and General Language Composite.
   (b) Freedom from Distractibility and Processing Speed Quotient.
   (c) General Memory Index and General Language Composite.
   (d) General Language Composite and Working Memory Index.

Answers: 1. True; 2. c; 3. Public Law 94-142, Education for All Handicapped Children Act of 1975; 4. False; 5. c; 6. True; 7. a.