One

OVERVIEW

The Wechsler Preschool and Primary Scale of Intelligence–Fourth Edition (WPPSI-IV; Wechsler, 2012) continues the progressive trend of recent Wechsler intelligence scale revisions that mirror contemporary advances in intelligence theory, neuropsychology, cognitive neuroscience, and psychometric methodology. Major modifications have been made to both the content and structure to reflect these advances. Primary index scores provide reliable and valid estimates of ability in several distinct but related areas of cognitive functioning, including verbal comprehension, visual-spatial ability, fluid reasoning, working memory, and processing speed. A Full Scale IQ score also is available to represent the child's overall level of ability across these cognitive domains. A number of new, ancillary index scores (e.g., Vocabulary Acquisition Index, Nonverbal Index) are available to represent vocabulary acquisition, global intellectual ability, and cognitive proficiency in more specific clinical situations, such as referrals for suspected language delays, preliteracy concerns, school readiness, or evaluations of children who may have limited levels of English language fluency.

WPPSI-IV scores can be interpreted from both a normative and an intrapersonal perspective. They most often are used from a normative perspective; that is, to describe a child's cognitive ability by comparing the child's scores to those obtained by children of approximately the same age (i.e., comparison to a normative reference group). A child's intrapersonal pattern of cognitive strengths and weaknesses can be evaluated by comparing scores to an overall indicator of performance, or to one another. The score comparison approach has been reorganized and expanded for the WPPSI-IV, with increased comparison score options (e.g., the mean primary index score, the Full Scale IQ, the mean scaled score for the primary index subtests). When combined with the availability of numerous theoretically and practically based index scores (termed *ancillary* index scores in the published test), the WPPSI-IV interpretive approach allows a thorough evaluation of a child's pattern of cognitive strengths and weaknesses that can be tailored to the unique clinical situation each child presents.

Similar to authors of previous books within the *Essentials* series (e.g., Flanagan & Kaufman, 2009; Lichtenberger & Kaufman, 2004, 2013), our goal for this book is to provide a go-to reference for both novice and proficient practitioners using the WPPSI-IV. Administration, scoring, and interpretive information is clearly and succinctly covered in successive chapters, incorporating the familiar Rapid Reference, Caution, and Don't Forget boxes that are hallmark features of the *Essentials* series. We also include Behind the Scenes boxes that offer helpful insights into the test development process as we were the WPPSI-IV research directors. Test questions are included at the conclusion of each chapter to highlight critical content.

The CD included with this book contains appendix matter, such as administration aids, interpretive tables, and normative data for a number of additional index scores not available within the published test. These additional index scores were developed to meet interpretive needs stemming from varied practical and theoretical perspectives (e.g., Cattell-Horn-Carroll [CHC] and neuropsychological). The CD also includes the WPPSI-IV Interpretive Assistant 1.0, scoring software that calculates norms for the additional index scores and walks the practitioner through our interpretive approach (see Chapter 4), including numerous score comparisons not available in the published test that can more fully inform interpretation.

HISTORICAL FOUNDATIONS OF EARLY CHILDHOOD ASSESSMENT

Assessment of preschoolers and young children is a unique and specialized endeavor, different from that experienced with older children, adolescents, and adults. Similarly, some of the historical foundations for early childhood assessment are shared with school-age children and adults, whereas others are unique to young children of preschool age. The following sections include a brief history of key scientific and societal influences on early childhood assessment. For more comprehensive coverage of these topics, the reader is referred to Ford, Kozey, and Negreiros (2012); Kelley and Surbeck (2007); and Wortham (2012).

The recognition of childhood as a unique stage in the life cycle was a critical precursor for increasing our understanding of children's growth and development. Although early publications emphasized this difference and the importance of rearing and educating children (e.g., Locke, 1692; Rousseau, 1762), systematic efforts to study the cognitive development of young children did not begin until the latter part of the 19th century, when increased societal attention in European countries was directed toward the health and welfare of young children. This focus spawned an era referred to as the Child Study Movement, in which attempts were made to apply the scientific method to the study of children. G. Stanley Hall and Lawrence Frank played pivotal roles in bringing this movement to the United States, with Hall establishing the first major child study center at Clark University in Massachusetts in 1893. Frank further entrenched the movement in educational institutions by establishing numerous child study centers at universities across the United States, with funding provided by the Laura Spelman Rockefeller Memorial (Wortham, 2012).

The Child Study Movement produced considerable knowledge and understanding of social, emotional, and cognitive childhood development, but the program's reliance on primarily observational data from children in a wide variety of group settings limited its acceptance by psychologists and other members of the scientific community. Despite the movement's lost momentum in the early 20th century, Hall's students, including Lewis Terman, Arnold Gesell, and John Dewey, continued to serve as pioneers in the more scientifically based field of child psychology that was beginning to take root in the United States. Terman was instrumental in adapting the Binet measures for extensive use in the United States, and Gesell defined and described the characteristic behaviors arising at specific periods during early childhood development (Gesell & Amatruda, 1941). Dewey focused on educational reform and improving educational programs aimed at young children (Wortham, 2012).

Concurrent with the Child Study Movement, progress in the study of childhood intelligence was occurring in the standardized testing field. The influx of students with diverse backgrounds that resulted from compulsory schooling in France and the United States produced a need for a method to classify children for proper educational placement. Alfred Binet and Theodore Simon were among those commissioned to devise a means for identifying children who were unlikely to benefit from formal education in the Paris schools. Binet and Simon produced their first 30-item intelligence scale for this purpose in 1905. Items were ordered by increasing difficulty, and were scored using a pass/fail criterion. Although somewhat crude psychometrically relative to today's standards, the original Binet-Simon scale and its subsequent revisions (Binet, 1911; Binet & Simon, 1908) represented major advances over the obsolete, sensory-based intelligence measures, and incorporated mental tasks that measured reasoning, comprehension, and judgment (Boake, 2002). The Binet scales emphasized the need for a standard administration, simple scoring, and evidence that a test serves its intended purpose (e.g., to identify children with low cognitive ability).

Multiple translations of the Binet scales were completed in the United States, including a well-known version by Goddard and colleagues at the Vineland

Training School (Goddard, 1916) and one by Kuhlmann (1912) that attempted to lower the age range to 2 months by including additional items. Shortly before World War I (in 1916), Terman revised the Binet scale at Stanford. His revision expanded the age range upward to adulthood and replaced the mental age score with the intelligence quotient (IQ). The 1937 Terman-Merrill revision of the Stanford-Binet included a number of new preschool-level items as well as additional nonverbal and memory items.

Intelligence tests for young children that were available in the early part of the 20th century generally failed to capture the complexities and uniqueness of early childhood intelligence, and instead focused on the developmental trajectories of mental and physical skills (Lichtenberger & Kaufman, 2004). A number of intelligence tests for infants and preschoolers published in the middle of the 20th century, including the Cattell Infant Intelligence Scale (Cattell, 1940), the Full Range Picture Vocabulary Test (Ammons & Ammons, 1948), and the Leiter International Performance Scale (Leiter, 1948). Despite increasing options, Terman's adaptation and translation of Binet and Simon's intelligence test, the Stanford-Binet Intelligence Scales (1937) remained entrenched as the preferred intelligence measure in the United States for children until the latter part of the 20th century (Lichtenberger & Kaufman, 2004).

Prior to the emergence of the Child Study Movement and the rise of standardized testing, prevalent views of intelligence posited that it was a genetically determined, immutable trait that was manifested behaviorally through sensory functions (Kelley & Surbeck, 2007). Binet was one of the first to question this basic assumption, but the heredity versus environment debate would continue for more than 40 years, when the preponderance of evidence suggested an environmental role in cognitive development. The influential writings of Piaget (1952) were indicative of this fundamental shift in thought, due to his emphasis on the interaction between the child and his or her environment in shaping progression through the mental developmental stages. In line with Piaget's theory and mounting evidence in support of the environment's role in cognitive development, educators had also noted a persistent pattern of lower performance in children from poor backgrounds (Kelley & Surbeck, 2007). Thus, a renewed focus on the health and welfare of young children ensued, culminating in new social programs and educational legislation that included the Head Start program.

Head Start was the primary federal program established to improve the academic performance of children from economically disadvantaged homes, as well as English language learners. A wide variety of Head Start programs were established throughout the United States, ranging from traditional nursery schools to highly structured academic settings; however, all federally funded programs were required to provide proof of their effectiveness. This requirement resulted in a tremendous increase in the number of available instruments, with more than 200 childhood measures published between 1960 and 1980 (Kelley & Surbeck, 2007; Wortham, 2012). Publications of comprehensive intelligence measures for young children during this time period included the McCarthy Scales of Children's Abilities (McCarthy, 1972), and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967). Relative to their predecessors, the available instruments for preschoolers were improved, but not as psychometrically sound or as developmentally appropriate as those for older children and adults.

The Education for All Handicapped Children Act of 1975 (PL 94-142) made a profound impact on early childhood assessment by requiring that a free public education be provided to all children aged 3–21, regardless of handicap. The law also stipulated that children with handicaps between the ages of 6 and 21 be placed in the least restrictive educational environment based on placement decisions derived from nondiscriminatory evaluations. Similar recommendations were made, but not required, for children between the ages of 3 and 5. Subsequent amendments passed in 1986 (PL 99-457) addressed this shortcoming through creation of the Federal School Program, which extended the rights to children with disabilities between the ages of 3 and 5 (Wortham, 2012). The 1986 amendments also offered financial incentives to states for the provision of early childhood intervention programs for children from birth to 3 years of age (Ford et al., 2012).

More recent legislation such as the Americans with Disabilities Act (1990) further defined the requirements of special education and added the categories of autism and traumatic brain injury. Goals of the legislation included maximal inclusion and integration of children with disabilities into all educational areas, and the requirement that all childhood programs be prepared to serve children with disabilities (Wortham, 2012). In 1997, Congress reauthorized the Education for All Children Act of 1975, requiring states to include special education students in the yearly state testing and to publicly report the results. The No Child Left Behind Act of 2001 (NCLB) and the Individuals with Disabilities Improvement Act of 2004 further specified the requirement by noting that 95% of children with disabilities must participate in the testing. The effects of this legislation and the resulting initiatives have yet to be evaluated to determine if the needs of these children are being met in the intended manner. Regardless of the findings, the importance of developing reliable and valid measures of mental abilities for children with and without disabilities has received continued support over the past 50 years from changes in social policy and educational legislation.

Today, there are numerous measures of intellectual or cognitive ability available for use with preschoolers and young children. Rapid Reference 1.1 lists the most

Rapid Reference 1.1

21st Century Preschool Intelligence Measures

Measure	Common Abbreviation	Publication Information	Age Range	Primary Scores
Wechsler Preschool and Primary Scale of Intelligence–Fourth Edition	WPPSI-IV	Wechsler, 2012: Pearson	2:6–3:11 (Younger Battery) 4:0–7:7 (Older Battery)	Full Scale IQ Verbal Comprehension Index Visual Spatial Index Fluid Reasoning Index Working Memory Index Processing Speed Index Verbal Acquisition Index Nonverbal Index General Ability Index Cognitive Proficiency Index
Differential Ability Scales–Second Edition	DAS-II	Elliott, 2007: Pearson	2:5–3:5 (Early Years Lower Level) 3:6–6:11 (Early Years Upper Level)	General Conceptual Ability Special Nonverbal Composite Verbal Nonverbal Reasoning Spatial Working Memory Processing Speed School Readiness

Measure	Common Abbreviation	Publication Information	Age Range	Primary Scores
Woodcock–Johnson Tests of Cognitive Ability–Third Edition Normative Update	WJ III COG NU	Woodcock, McGrew, & Mather, 2007: Riverside	2:0–5:11 (Early Development Battery)	General Intellectual Ability Gc: crystallized knowledge Gv: visual–spatial ability Gf: fluid reasoning Gsm: short-term memory Gq: quantitative knowledge
Stanford–Binet Intelligence Scales for Early Childhood, Fifth Edition	Early SB5	Roid, 2005: Pro-Ed	2:0–7:3 (Early SB5)	Full Scale IQ Abbreviated Battery IQ Verbal IQ Nonverbal IQ Knowledge Visual–Spatial Processing Fluid Reasoning Working Memory Quantitative Reasoning
Kaufman Assessment Battery for Children–Second Edition	KABC-II	A. S. Kaufman & Kaufman, 2004: Pearson	3:0–3:11 (5 Subtests) 4:0–6:11 (10 Subtests)	Mental Processing Index Fluid–Crystallized Index Knowledge Nonverbal Index Sequential Simultaneous Learning
Reynolds Intellectual Assessment Scales	RIAS	Reynolds & Kamphaus, 2003:	3:0–94 (All Subtests)	Composite Intelligence Index

(*continued*)

Measure	Common Abbreviation	Publication Information	Age Range	Primary Scores
		Psychological Assessment Resources		Verbal Intelligence Index Nonverbal Intelligence Index Composite Memory Index

commonly used measures that have published since 2000, as well as some basic descriptive information. A thorough review of these measures is beyond the scope of this chapter; however, Ford et al. (2012) include a concise summary of each scale's strengths and weaknesses.

HISTORICAL FOUNDATIONS OF THE WPPSI-IV

Excellent accounts of the historical foundations of intelligence testing have been written (e.g., Goldstein & Beers, 2003; R. J. Sternberg, 2000; Wasserman, 2012), and the reader is referred to these sources to gain a greater understanding of historical developments in this area. As with most texts on the Wechsler intelligence scales, we have elected to include a brief section describing the historical foundations of the Wechsler scales, despite the possibility that the inclusion of such information may serve to perpetuate a misperception that the scales are outdated. It is our contention that Wechsler's foresight to define intelligence in practical terms allowed the necessary flexibility for continuous revisions to the scales in light of advances in theory, research, and the measurement of intelligence (Coalson, Raiford, Saklofske, & Weiss, 2010). His 1939 definition of intelligence is still relevant today and continues to appear in the most recent revisions of his scales (Wechsler, 2003, 2008, 2012). Wechsler defined intelligence as

> the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his [or her] environment. It is global because it characterizes the individual's behavior as a whole; it is an aggregate because it is composed of elements or abilities which, though not entirely independent, are qualitatively differentiable.
>
> *(Wechsler, 1944, p. 3)*

Wechsler thus supported the existence of general or global intelligence, but also acknowledged that general intelligence is composed of qualitatively different abilities (e.g., verbal comprehension, visual–perceptual skills, and reasoning ability). He believed that intelligence was more than just cognitive abilities and that nonintellective, conative factors such as curiosity, drive, and persistence contributed to the expression of intelligence (Wechsler, 1950). His astute clinical skills were evident in his selection of subtests for his batteries, all of which have been shown to measure important factors of intelligence since their introduction (Carroll, 1993, 2012; Horn & Blankson, 2012; W. J. Schneider & McGrew, 2012).

Consistent with Wechsler's definition of intelligence, results of comprehensive factor-analytic investigations of cognitive ability measures suggest overwhelming evidence for a general intelligence factor at the apex of a hierarchical construct that is composed of a set of related but distinguishable cognitive abilities (Carroll, 1993, 2012). Intelligence appears to be composed of 8 to 10 broad domains that are, in turn, composed of more specific abilities (Carroll, 1993; Horn & Noll, 1997). Additional research is needed to determine whether all of these domains are present in young children. Although some research suggests that the number of intelligence factors is reduced in young children relative to older children and adults (Morgan, Rothlisberg, McIntosh, & Hunt, 2009; Ward, Rothlisberg, McIntosh, & Bradley, 2011), other studies suggest that more differentiation among young children's cognitive abilities exists than was once believed (e.g., S. B. Kaufman, Reynolds, Liu, Kaufman, & McGrew, 2012; Kuwajima & Sawaguchi, 2010; Morgan et al., 2009; W. Schneider, Schumann-Hengsteler, & Sodian, 2005).

As indicated in Rapid Reference 1.2, the WPPSI-IV retains a number of subtests with origins in 20th century measures, supplying additional evidence of Wechsler's impressive clinical judgment when selecting the subtests to include in his original scales.

Extensive testing of military recruits during World War I raised questions about the limitations and weaknesses of early intelligence tests. In particular, concerns with the validity of IQ scores across the age range and for specific ethnic and socioeconomic groups were noted, as well as the increasing factor-analytic evidence suggesting that intelligence was composed of several abilities. Such was the atmosphere when David Wechsler entered the field of test development. Wechsler's clinical experiences as a psychological examiner for military recruits and his psychometric training under Charles Spearman and Karl Pearson led him to develop an adult intelligence scale that addressed some of the shortcomings he had personally observed (Boake, 2002).

Rapid Reference 1.2

Origins of WPPSI-IV Subtests

Subtest	Origin	First WPPSI Edition
Information	Army Alpha Group Examination	WPPSI
Similarities	*Stanford-Binet*	WPPSI
Vocabulary	*Stanford-Binet*	WPPSI
Comprehension	*Stanford-Binet*	WPPSI
	Army Alpha Group Examination	
Receptive Vocabulary	*Stanford-Binet*	WPPSI-III
Picture Naming	*Stanford-Binet*	WPPSI-III
Block Design	*Army Performance Scale*	WPPSI
	Kohs Block Design (1923)	
Object Assembly	*Pintner-Paterson performance tests* (1917)	WPPSI-R
	Army Performance Scale	
Matrix Reasoning	*Raven's Progressive Matrices* (1938)	WPPSI-III
Picture Concepts	Novel task developed by Pearson	WPPSI-III
Picture Memory	Novel task developed by Pearson	WPPSI-IV
Zoo Locations	Novel task developed by Pearson	WPPSI-IV
Bug Search	Schneider and Shiffrin (1977)	WPPSI-IV
	Sternberg (1966)	
Cancellation	Albert (1973)	WPPSI-IV
	WISC-IV Cancellation	
Animal Coding	*Substitution Test* (Kirkpatrick, 1909)	WPPSI-IV

His first test, the Wechsler-Bellevue Intelligence Scale (Wechsler, 1939) possessed important advantages relative to its competitors, including both verbal and nonverbal measures in a single test. More importantly, Wechsler introduced the use of deviation IQs, which offered increased score comparability across measures and improved accuracy, interpretability, and clinical utility. Wechsler subsequently published a downward extension of the Wechsler-Bellevue scale as the Wechsler Intelligence Scale for Children (WISC, 1949), which was designed for children ages 5 to 15. There were few changes from the adult form of the scale, although easier items had been added to extend the test's floor. The Wechsler-Bellevue was revised again in 1955 and published as the Wechsler Adult Intelligence Scale (WAIS). In response to the increasing societal and clinical needs for measures of early childhood intelligence, the WPPSI (Wechsler, 1967) was developed for children between the ages of 4 years 0 months and 6 years 6 months. The Wechsler intelligence scales rose to a level of prominence among assessment measures, with continuous revisions spanning over 70 years.

The first revision of the WPPSI (WPPSI-R; Wechsler, 1989) extended the age range to children between the ages of 3 years 0 months and 7 years 3 months, and retained all of the original WPPSI subtests (Information, Comprehension, Arithmetic, Vocabulary, Similarities, Sentences, Geometric Design, Block Design, Mazes, Picture Completion, and Animal House [renamed Animal Pegs]). Easier and more difficult items were developed to extend the floors and ceilings of several retained subtests, and Object Assembly was added. Although children were administered every subtest in both WPPSI and WPPSI-R, age-determined start points were introduced to reduce administration time. Subsequent factor-analytic studies supported a two-factor model of intelligence for both scales, including both Verbal and Performance factors (Carlson & Reynolds, 1981; Gyurke, Stone, & Beyer, 1990; B. Schneider & Gervais, 1991; Silverstein, 1969; Stone, Gridley, & Gyurke, 1991).

The subsequent revision of the WPPSI-R resulted in the *Wechsler Preschool and Primary Scale of Intelligence–Third Edition* (WPPSI-III; Wechsler, 2002), a scale with more dramatic differences from its predecessor than those in the previous revision. The most obvious change in the WPPSI-III was the division of the covered age range into two age bands with different subtest batteries; ages 2:6 to 3:11 and ages 4:0 to 7:3. Five WPPSI-R subtests were dropped for the WPPSI-III, including Arithmetic, Animal Pegs, Geometric Design, Mazes, and Sentences. These subtests were replaced with five new subtests that were more developmentally appropriate and designed to measure constructs that were shown to be important aspects of intelligence. Similar to previous revisions of the WISC and WAIS, the Coding and Symbol Search subtests were added to the WPPSI-III as measures of processing speed. Word Reasoning, Matrix Reasoning, and Picture Concepts were added as new measures of fluid reasoning.

Revisions to the Wechsler intelligence scales are based on psychometric and theoretical advances, as well as clinical research and practical need: They are not based on fundamental changes to Wechsler's definition of intelligence or any single theory of intelligence or cognitive development. Regrettably, we never had the opportunity to meet or work with David Wechsler. However, based on his accomplishments and writings, we believe he would have embraced advancements in the development of his instruments, based on guidance from contemporary theories of intelligence, child development, neuroscience, and other related fields.

We are confident he would have insisted on the abundant evidence of psychometric quality and clinical utility. The changes in the WPPSI-IV continue to reflect this revision trend and are detailed in subsequent sections of this chapter.

DEVELOPMENT OF THE WPPSI-IV

Key Revisions

A variety of issues precipitated the WPPSI-IV revision. The *WPPSI-IV Technical and Interpretive Manual* (*Technical and Interpretive Manual*; Wechsler, 2012) discusses these issues in detail on pages 19–31. Rapid Reference 1.3 lists the key revision features broadly and specifically.

Rapid Reference 1.3

Broad and Detailed Key Revisions

Broad Key Revision	Detailed Key Revisions
Updated theoretical foundations	• Incorporate and consider research on contemporary structural intelligence models • Incorporate and consider neurodevelopmental and neurocognitive research • Incorporate and consider working memory models and research
Increased developmental appropriateness	• Improve the developmental appropriateness of manipulatives • Improve the developmental appropriateness of instructions • Improve the developmental appropriateness of the Processing Speed subtests
Increased user friendliness	• Enhance item security • Improve user friendliness of materials and packaging • Minimize testing time • Improve user friendliness of administration and scoring • Reduce length of discontinue rules
Improved psychometric properties	• Improve psychometric properties of items and scoring rules • Update the norming method • Increase evidence of reliability and validity

Broad Key Revision	Detailed Key Revisions
	• Improve subtest floors and ceilings • Reduce item bias • Expand critical value significance level options
Enhance clinical utility	• Improve the clinical utility of the test structure • Organize the score differences comparison methodology to maximize clinical utility • Extend the age range upward • Reduce the expressive language requirements necessary to obtain a composite score • Provide ancillary index scores with specific clinical applications • Increase the number of special group studies • Provide statistical linkage to a measure of achievement, and build in a pattern of strengths and weaknesses analysis

Subtests

Practitioners who used the WPPSI-III will find many of the same core Wechsler subtests are present but substantively revised (with many new items and in some cases new procedures). They also will notice a number of new subtests, a modified test structure, more composite scores, and a new approach to score analysis on the Record Form.

New Subtests

There are five new subtests:

1. Picture Memory, a visual Working Memory subtest that utilizes proactive interference rather than sequencing to introduce cognitive processing demands necessary for measuring working memory.
2. Zoo Locations, a visual-spatial Working Memory subtest that also relies on proactive interference to introduce cognitive processing demands.
3. Bug Search, a timed visual matching Processing Speed subtest inspired by WPPSI-III Symbol Search.
4. Cancellation, a speeded visual search Processing Speed subtest inspired by WISC-IV (Wechsler, 2003) Cancellation.
5. Animal Coding, a timed visual paired associates Processing Speed subtest inspired by WPPSI-III Coding and WPPSI-R Animal Pegs.

Information about the development of these new subtests that provides insight into the test development process appears in the Behind the Scenes boxes in Chapter 2 of this book.

Dropped Subtests

Four subtests were removed from the WPPSI-III complement prior to WPPSI-IV development. These subtests were removed for varying reasons:

- Word Reasoning was dropped because it conceptually overlapped with the Vocabulary subtest, it lacked evidence of measuring fluid reasoning (which was the original intent), and it was strongly associated with the Information subtest to the point of psychometric redundancy. Furthermore, it was not well liked by children and not rated highly by practitioners for its clinical utility. It was retained for the prior edition because of its superior floors relative to Similarities and Comprehension, which were substantially improved for this revision. Thus, it was no longer necessary.
- Picture Completion was deleted because it was desirable to reduce the emphasis of the test on speeded performance and fine visual detail discrimination, and to make room for the Working Memory subtests.
- Symbol Search and Coding were removed to make room for the new Processing Speed subtests, which measure similar constructs in that domain but are more developmentally appropriate and reduce greatly the reliance on fine motor skills. Notably, both of these subtests had a naturally occurring floor that couldn't be overcome without redesigning the tasks; hence, the new subtests were developed.

Retained Subtests

For ages 2:6 to 3:11, five WPPSI-III subtests were retained. For ages 4:0 to 7:7, 10 subtests were retained. Rapid Reference 1.4 lists the retained subtests by age band and examples of changes made to those subtests. The revisions are more specifically detailed in Chapters 2 and 3 of this book.

Subtest Descriptions and Expert References on Constructs Measured and Abilities Engaged

Rapid Reference 1.5 provides a description of all subtests, reproduced by permission from the test publisher. New subtests are indicated with an asterisk. The age range for each subtest is also listed, as not all subtests are available for children aged 2:6 to 3:11. Rapid Reference 1.6 provides information on the constructs and abilities ascribed to each subtest.

Rapid Reference 1.4

Retained Subtests by Age Band and Changes

Subtest	Ages	Changes
Information	2:6–7:7	New and revised items and scoring criteria
		Updated with more child-appropriate and contemporary questions
		Reduced total items and shorter discontinue rule
Similarities	4:0–7:7	New and revised items and scoring criteria
		Introduced picture items to improve the subtest floor
		Reduced total items and shorter discontinue rule
Vocabulary	4:0–7:7	New and revised items and scoring criteria
		Reduced total items and shorter discontinue rule
Comprehension	4:0–7:7	New and revised items and scoring criteria
		Introduced picture items to improve the subtest floor
		Updated with more child-appropriate and contemporary questions
		Shorter discontinue rule
Receptive Vocabulary	2:6–7:7	New and revised items
		Reduced total items and shorter discontinue rule
Picture Naming	2:6–7:7	New and revised items
		Reduced total items and shorter discontinue rule
Block Design	2:6–7:7	New and revised items
		Increased teaching and transition between one-color and two-color blocks
		New items to extend the ceiling
		Reduced total items and shorter discontinue rule
Object Assembly	2:6–7:7	New and revised items
		New item to extend the subtest ceiling
		Reduced total items and shorter discontinue rule
Matrix Reasoning	4:0–7:7	New and revised items
		Reduced total items and shorter discontinue rule
Picture Concepts	4:0–7:7	New and revised items
		Reduced total items and shorter discontinue rule

Rapid Reference 1.5

Subtest Abbreviations, Descriptions, and Age Ranges

Subtest	Abbreviation	Description	Age Range
Information	IN	For picture items, the child selects the response option that best answers a question about a general-knowledge topic. For verbal items, the child answers questions about a broad range of general-knowledge topics.	2:6–7:7
Similarities	SI	For picture items, the child selects the response option that is from the same category as two other depicted objects. For verbal items, the child is read two words that represent common objects or concepts and describes how they are similar.	4:0–7:7
Vocabulary	VC	For picture items, the child names the depicted object. For verbal items, the child defines words that are read aloud.	4:0–7:7
Comprehension	CO	For picture items, the child selects the response option that represents the best response to a general principle or social situation. For verbal items, the child answers questions based on his or her understanding of general principles and social situations.	4:0–7:7
Receptive Vocabulary	RV	The child selects the response option that best represents the word the examiner reads aloud.	2:6–7:7
Picture Naming	PN	The child names depicted objects.	2:6–7:7
Block Design	BD	Working within a specified time limit, the child views a model and/or a picture and uses one- or two-color blocks to re-create the design.	2:6–7:7

Subtest	Abbreviation	Description	Age Range
Object Assembly	OA	Working within a specified time limit, the child assembles the pieces of a puzzle to create a representation of an identified object.	2:6–7:7
Matrix Reasoning	MR	The child views an incomplete matrix and selects the response option that completes the matrix.	4:0–7:7
Picture Concepts	PC	The child views two or three rows of pictures and selects one picture from each row to form a group with a common characteristic.	4:0–7:7
Picture Memory*	PM	The child views a stimulus page of one or more pictures for a specified time and then selects the pictures from options on a response page.	2:6–7:7
Zoo Locations*	ZL	The child views one or more animal cards placed on a zoo layout for a specified time and then places each card in the previously viewed locations.	2:6–7:7
Bug Search*	BS	Working within a specified time limit, the child marks the bug in the search group that matches the target bug.	4:0–7:7
Cancellation*	CA	Working within a specified time limit, the child scans two arrangements of objects (one random, one structured) and marks target objects.	4:0–7:7
Animal Coding*	AC	Working within a specified time limit and using a key, the child marks shapes that correspond to pictured animals.	4:0–7:7

* New subtest

Source: Table 1.1 of the WPPSI-IV Administration and Scoring Manual.

Rapid Reference 1.6

WPPSI-IV Subtest Constructs and Abilities

Subtest	Rationale
Information	Designed to measure: Acquisition, retention, and retrieval of general facts
	Related to: Crystallized ability, Gc-K0 (general information), and retention and retrieval of learned information, Glr
	May also involve: Auditory perception, verbal expression
Similarities	Designed to measure: Verbal concept formation and abstract reasoning
	Related to: Crystallized ability, associative and categorical thinking, Gf-I (induction), concept recognition and generation
	May also involve: Auditory perception and verbal expression
Vocabulary	Designed to measure: Word knowledge, verbal concept formation
	Related to: Crystallized ability, Gc-VL (lexical knowledge), fund of knowledge, learning, verbal expression, long-term memory
	May also involve: Auditory perception, auditory comprehension, abstract thinking, expressive vocabulary
Comprehension	Designed to measure: Verbal reasoning, verbal conceptualization, verbal comprehension, verbal expression, practical knowledge, judgment
	Related to: Crystallized ability (Gc), understanding of societal standards and conventional behavior, social judgment, Glr, common sense
	May also involve: Auditory perception
Receptive Vocabulary	Designed to measure: Word knowledge, verbal concept formation, receptive vocabulary
	Related to: Crystallized ability, Gc-VL (lexical knowledge), fund of knowledge, learning, verbal expression, long-term memory
	May also involve: Visual perception, auditory comprehension
Picture Naming	Designed to measure: Word knowledge, verbal concept formation, expressive vocabulary
	Related to: Crystallized ability, Gc-VL (lexical knowledge), fund of knowledge, learning, verbal expression, long-term memory
	May also involve: Visual and auditory perception

Subtest	Rationale
Block Design	Designed to measure: Visual–spatial processing, analysis and synthesis of abstract visual stimuli Related to: Gv-SR (spatial relations), Gv-Vz (visualization), Gv-CS (closure speed), mental rotation, nonverbal reasoning, visual perception, simultaneous processing, problem solving, cognitive flexibility, planning May also involve: Visual–motor coordination
Object Assembly	Designed to measure: Visual–spatial processing, analysis and synthesis of meaningful visual stimuli Related to: Gv-SR, Gv-Vz, Gv-CS, mental rotation, nonverbal reasoning, visual perception, simultaneous processing, problem solving, cognitive flexibility, planning May also involve: Visual–motor coordination
Matrix Reasoning	Designed to measure: Fluid reasoning/intelligence, classification ability, inductive reasoning Related to: Gf-I (induction), Gf-RG (general sequential reasoning), simultaneous and successive processing, planning, metacognition, problem solving, cognitive flexibility, reasoning, planning May also involve: Visual perception
Picture Concepts	Designed to measure: Fluid reasoning/intelligence, classification ability, inductive reasoning Related to: Gf-I, Gf-RG, simultaneous and successive processing, planning, metacognition, concept recognition and generation, problem solving, cognitive flexibility, reasoning, planning May also involve: Gc (crystallized ability), acquired knowledge
Picture Memory	Designed to measure: Visual working memory, ability to withstand proactive interference Related to: Gsm-MW (working memory capacity), Gsm-MS (memory span), Gv-MV (visual memory), attention, simultaneous and successive processing, planning and metacognition, visual immediate memory for pictures, response inhibition May also involve: Visual perception
Zoo Locations	Designed to measure: Visual–spatial working memory, ability to withstand proactive interference Related to: Gsm-MW (working memory capacity), Gsm-MS (memory span), Gv-MV (visual memory), attention, simultaneous and successive processing, planning and metacognition, visual immediate memory for pictures and spatial locations, response inhibition May also involve: Visual perception, visual–motor construction

(*continued*)

Subtest	Rationale
Bug Search	Designed to measure: Processing speed, simple visual discrimination Related to: Gs-P (perceptual speed), Gs-R9 (rate of test taking), simultaneous processing, planning and metacognition, speed and efficiency, selective and sustained attention, visual scanning and tracking, visual immediate memory for pictures, response inhibition May also involve: Visual–motor skills
Cancellation	Designed to measure: Processing speed Related to: Gs-P (perceptual speed), Gs-R9 (rate of test taking), simultaneous processing, planning and metacognition, speed and efficiency, selective and sustained attention, visual scanning and tracking, visual immediate memory for pictures, response inhibition, classification ability May also involve: Visual–motor skills, acquired knowledge
Animal Coding	Designed to measure: Processing speed Related to: Gs-P (perceptual speed), Gs-R9 (rate of test taking), simultaneous processing, planning and metacognition, speed and efficiency, selective and sustained attention, visual scanning and tracking, visual immediate memory for pictures and objects, response inhibition, visual associative memory May also involve: Visual–motor skills

Note. References: Carroll, 1993; Flanagan, Alfonso, and Ortiz, 2012; Flanagan, Alfonso, Ortiz, and Dynda, 2010; Groth-Marnat, 2003; Lichtenberger and Kaufman, 2004, 2013; Miller, 2010, 2013; Sattler, 2008; W. J. Schneider and McGrew, 2012.

Subtest Terminology

The WPPSI-IV categorizes subtests into three categories—core, supplemental, and optional—that indicate a subtest's status in relation to a given composite score. The categorical assignment sometimes differs across the two age bands (i.e., 2:6–3:11 and 4:0–7:7) as well as across different composite scores. Furthermore, unlike the WISC-IV and the WAIS-IV (Wechsler, 2008), subtests that are core for a given index score are not necessarily core for the Full Scale IQ, because the Full Scale IQ is not derived from every subtest that is a core index subtest.

Core subtests are used to derive the composite score normative information and values. Supplemental subtests are provided to allow assessment of additional intellectual ability constructs, and can be used in some situations as substitutes for core subtests when a necessary subtest score is missing. Optional subtests are included to complement the existing intellectual ability information, but cannot substitute for a missing or invalid core subtest.

Composite Scores

Composite Score Terminology

The 10 composite scores from the published test are described in detail in Chapter 4 of this book. Practitioners who used the WPPSI-III will notice important changes to the composite score terminology. Some scores were renamed, and new scores were created.

The most obvious change is that the names Verbal IQ and Performance IQ are no longer in use. This elimination began with the WISC-IV and WAIS-IV, and the terms are now completely phased out. The Verbal IQ was renamed as the Verbal Comprehension Index. For children aged 2:6 to 3:11, the Performance IQ is renamed the Visual Spatial Index; for those ages it is derived from the same subtests that contributed to that score. However, for children 4:0 to 7:7, the Performance IQ as it stood no longer exists. In its place for these ages there are now two scores, the Visual Spatial Index and the Fluid Reasoning Index. The WPPSI-III Performance IQ formerly was derived from subtests that now contribute to these two separate index scores. The WPPSI-III Processing Speed Quotient was renamed the Processing Speed Index, as was done previously on the WISC-IV and the WAIS-IV, to improve consistency across the different scales. Finally, the WPPSI-III General Language Composite is now termed the Vocabulary Acquisition Index to better represent the skills measured by its contributing subtests.

The published test divides the nine index scores into primary and ancillary categories. The primary index scores are thus termed because they are based on factor-analytic evidence and represent the main constructs measured within the test. The ancillary index scores are designed for use in specific practical and clinical situations.

The published composite scores are listed, with their abbreviations (used in some tables in this book and throughout the published manuals), in Rapid Reference 1.7. Because not all subtests are appropriate or available for ages 2:6 to 3:11, there are fewer composite scores available for younger children (seven). The age range for each composite score is therefore listed, along with the index score categorical membership (primary or ancillary).

Test Structure

The age range covered by the WPPSI-IV includes periods of great cognitive growth and development. For this reason, the battery is different for the two age bands (ages 2:6 to 3:11 and ages 4:0 to 7:7), and the composite scores for each age band are composed of different subtests.

Rapid Reference 1.7

Published Composite Score Abbreviations and Age Range

Composite Score	Abbreviation	Age Range	Index Score Category
Verbal Comprehension Index	VCI	2:6–7:7	Primary
Visual Spatial Index	VSI	2:6–7:7	Primary
Fluid Reasoning Index	FRI	4:0–7:7	Primary
Working Memory Index	WMI	2:6–7:7	Primary
Processing Speed Index	PSI	4:0–7:7	Primary
Full Scale IQ	FSIQ	2:6–7:7	n/a
Vocabulary Acquisition Index	VAI	2:6–7:7	Ancillary
Nonverbal Index	NVI	2:6–7:7	Ancillary
General Ability Index	GAI	2:6–7:7	Ancillary
Cognitive Proficiency Index	CPI	4:0–7:7	Ancillary

The 2:6 to 3:11 Age Band

For children aged 2:6 to 3:11, there are three primary index scores available (i.e., the Verbal Comprehension Index, Visual Spatial Index, and Working Memory Index), as well as a Full Scale IQ and three ancillary index scores (i.e., the Vocabulary Acquisition Index, Nonverbal Index, and General Ability Index). Each composite score is derived from the core subtests on the corresponding scale. If a subtest is listed on a scale in Figure 1.1, it does not indicate that it can automatically substitute for all core subtests on the scale. Refer to Chapter 3 of this book and to the *Administration and Scoring Manual* for in-depth discussion of the substitution, proration, and invalidation rules, which differ substantially from the WPPSI-III, WISC-IV, and WAIS-IV. Figure 1.1 depicts the test framework for 2:6 to 3:11. Subtests listed in bold font for a given scale are core to the corresponding composite score. Subtests that appear in italic font are supplemental for that scale. If no subtests are listed in italics, there are no supplemental subtests for that scale. Rapid Reference 1.1 provides the core subtest composition of the published composite scores, by age band.

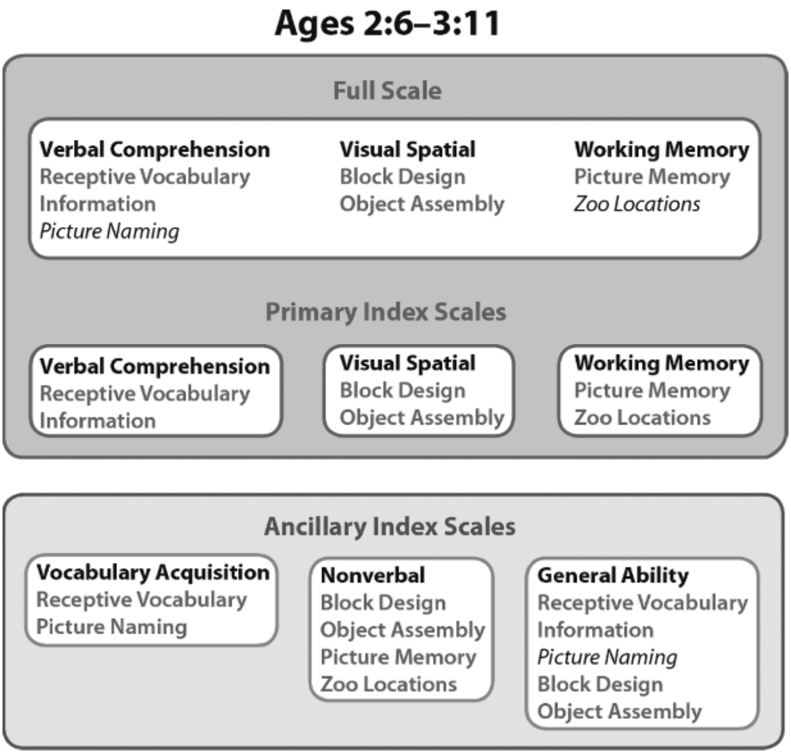

Figure 1.1 Test Structure for 2:6–3:11

The 4:0 to 7:7 Age Band

For children aged 4:0 to 7:7, there are five primary index scores available (i.e., the Verbal Comprehension Index, Visual Spatial Index, Fluid Reasoning Index, Working Memory Index, and Processing Speed Index), as well as a Full Scale IQ and four ancillary index scores (i.e., the Vocabulary Acquisition Index, Nonverbal Index, General Ability Index, and Cognitive Proficiency Index). Each composite score is derived from the core subtests on the corresponding scale. As with the younger age band, if a subtest is listed on a scale in Figure 1.2, it

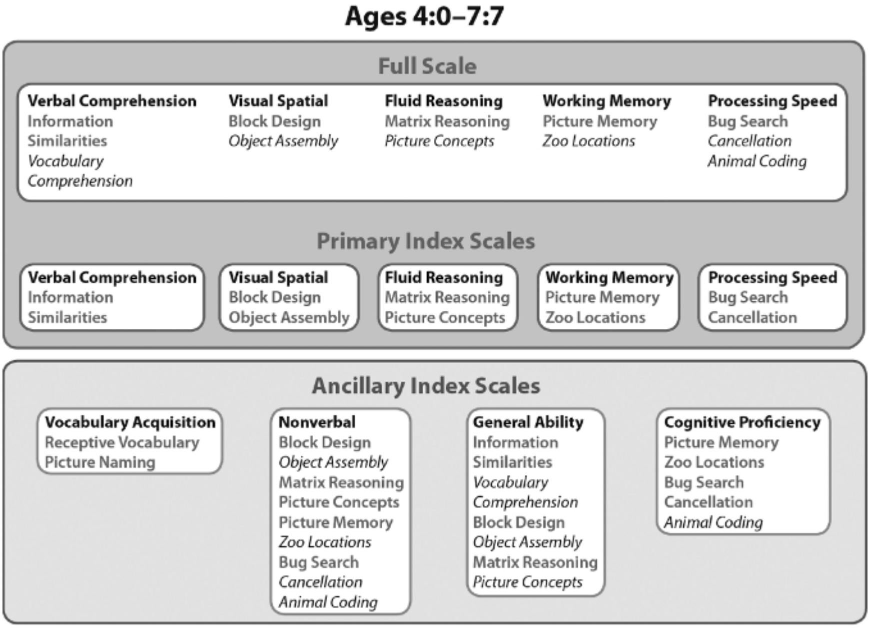

Figure 1.2 Test Structure for 4:0–7:7

does not indicate that it can automatically substitute for all core subtests on the scale. Refer to Chapter 3 of this book and to the *Administration and Scoring Manual* for in-depth discussion of the substitution, proration, and invalidation rules, which differ substantially from the WPPSI-III, WISC-IV, and WAIS-IV. Figure 1.2 depicts the test framework for 4:0 to 7:7. Subtests listed in bold font for a given scale are core to the corresponding composite score. Subtests that appear in italic font are supplemental for that scale. If no subtests are listed in italics, there are no supplemental subtests for that scale.

Subtest Composition of Published Composite Scores

Rapid Reference 1.8 lists the core subtests for the published composite scores for each age band.

Rapid Reference 1.8

Core Subtest Composition of Published Composite Scores, by Age Band

Subtest	VCI	VSI	FRI	WMI	PSI	FSIQ	VAI	NVI	GAI	CPI
IN	Y, O					Y, O			Y, O	
SI	O					O			O	
RV	Y					Y	Y, O		Y	
PN							Y, O			
BD		Y, O				Y, O		Y, O	Y, O	
OA		Y, O				Y		Y	Y	
MR			O			O		O	O	
PC			O					O		
PM				Y, O		Y, O		Y, O		O
ZL				Y, O				Y		O
BS					O	O		O		O
CA					O					O

Note. Y = core FSIQ subtest for ages 2:6–3:11, O = core FSIQ subtest for ages 4:0–7:7. Abbreviations are IN = Information, SI = Similarities, RV = Receptive Vocabulary, PN = Picture Naming, BD = Block Design, OA = Object Assembly, MR = Matrix Reasoning, PC = Picture Concepts, PM = Picture Memory, ZL = Zoo Locations, BS = Bug Search, CA = Cancellation, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, VAI = Vocabulary Acquisition Index, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.

Additional Index Scores in This Book and on the Accompanying CD

There are a number of additional index scores provided in this book and on the accompanying CD. The additional index scores were developed based upon specific theoretical approaches and practical considerations. The norms for these additional index scores are available on the CD that accompanies this book, which contains appendix matter and the WPPSI-IV Interpretive Assistant 1.0. Rapid Reference 1.9 provides a summary of the subtest composition of the additional index scores. Supporting technical evidence for these additional index scores is provided in Chapter 4 of this book.

Rapid Reference 1.9

Subtest Composition of Additional Index Scores

Subtest	Gc-K0	Gc-VL	Gf-Verbal	WKI	CRGI	CVI	CEI
IN	✓					✓	
SI			✓		✓	✓	
VC		✓		✓		✓	✓
CO	✓		✓			✓	✓
RV		✓					
PN		✓		✓			
PC					✓		

Note. IN = Information, SI = Similarities, VC = Vocabulary, CO = Comprehension, RV = Receptive Vocabulary, PN = Picture Naming, PC = Picture Concepts. Gc-K0 = Gc narrow ability of general information, Gc-VL = Gc narrow ability of lexical knowledge, Gf-Verbal = inductive reasoning with verbal stimuli, WKI = Word Knowledge Index, CRGI = Concept Recognition and Generation Index, CVI = Comprehensive Verbal Index, CEI = Complex Expressive Index.

VALIDITY

Factor Analytic Studies

As noted, the WPPSI-IV test structure represents a substantial revision relative to the WPPSI-III. With 4 subtests having been deleted and 5 added, the WPPSI-IV is composed of 7 subtests for 2:6 to 3:11 and 15 subtests for 4:0 to 7:7. As with the WISC-IV and the WAIS-IV before, the structure focuses on the primary-index scores as the primary level of interpretation because they are supported by factor analysis, as well as by clinical and practical utility.

The *Technical and Interpretive Manual* reports the results of several confirmatory factor analytic studies that each support the hierarchical three-factor structure for the younger age band, and the hierarchical five-factor structure for the older age band.

The studies were conducted by age band on two sets of subtests: first on all subtests, then on all primary index subtests. The younger age band results supported a hierarchical three-factor structure with the Full Scale at the apex and separate factors for Verbal Comprehension, Visual Spatial, and Working Memory, with each subtest loading on its expected factor and no cross loadings or correlated error permitted.

The older age band results indicated a hierarchical five-factor structure with the Full Scale at the apex and separate factors for Verbal Comprehension, Visual

Spatial, Fluid Reasoning, Working Memory, and Processing Speed. Each subtest loaded on its expected factor. The model fit was excellent, but examination of modification indices and residuals suggested that two nested subfactors were present within the Verbal Comprehension factor. One subfactor, labeled *Broad/Expressive*, contained Information, Similarities, Vocabulary, and Comprehension. The second subfactor, labeled *Focused/Simple*, contained Receptive Vocabulary and Picture Naming. The groupings were theoretically meaningful because the first contained subtests that required more expressive responses and were not limited to lexical knowledge, and the second contained subtests that required little to no expressive responses and were focused on word meanings only. The Focused/Simple factor provides some factor-analytic support for the Vocabulary Acquisition Index at these ages, and the Broad/Expressive subfactor supports permitting only Vocabulary or Comprehension to substitute for Information or Similarities within the older age band's Full Scale IQ.

The test structures for each age band were supported within both sets of subtests (i.e., all subtests, core index subtests only), with statistically significant improvement of fit relative to less complex models and excellent fit indices (e.g., .97 Tucker-Lewis Index for all subtests on each battery). The fit indices were subsequently run for the selected model to confirm the older age band results within more narrow age groups (i.e., 4:0 to 4:11, 5:0 to 5:11, and 6:0 to 7:7). Each produced excellent fit indices (e.g., .98, .97, and .95 Tucker-Lewis Index for all subtests for each of the narrow age groups).

Relation to WPPSI-III

The WPPSI-IV represents a rather dramatic departure from its predecessor. As indicated in this chapter, a number of the traditional Wechsler subtests remain, others have been replaced, new subtests and composite scores have been added that measure constructs not previous measured, and the factor structure for both age bands is quite different as a result. Furthermore, as discussed in Chapter 2 of this book, a large proportion of items were replaced on many subtests, and the administration and scoring rules were revised. Additionally, as discussed in greater detail in Chapter 4 of this book, new comparison score approaches have been added to the test.

Evaluating the relation of the current test to the prior edition informs judgments about applicability of the research base on the prior edition to the current, and about how results may differ on the new test. This evaluation is particularly relevant for the WPPSI-IV due to the aforementioned changes.

For comparison purposes, Rapid Reference 1.10 lists the composition of the WPPSI-III and the WPPSI-IV Full Scale IQ by age band.

Rapid Reference 1.10

Comparison of the WPPSI-III and WPPSI-IV Full Scale IQ, by Age Band

Subtest	WPPSI-III 2:6–3:11	WPPSI-IV 2:6–3:11	WPPSI-III 4:0–7:3	WPPSI-IV 4:0–7:7
IN	✓	✓	✓	✓
SI				✓
RV	✓	✓		
VC			✓	
WR			✓	
BD	✓	✓	✓	✓
OA	✓	✓		
MR			✓	✓
PC			✓	
PM		✓		✓
BS				✓
CD			✓	

Note. Y = Core for ages 2:6–3:11, O = core for ages 4:0–7:7.
Abbreviations are IN = Information, SI = Similarities, RV = Receptive Vocabulary, VC = Vocabulary, WR = Word Reasoning, BD = Block Design, OA = Object Assembly, MR = Matrix Reasoning, PC = Picture Concepts, PM = Picture Memory, BS = Bug Search, CD = Coding.

For children aged 2:6 to 3:11, the Full Scale IQ is largely unchanged. The most important change is the inclusion of a Working Memory subtest in the score. The remaining subtests that contribute to the Full Scale IQ are the same as for the WPPSI-III. Hence, the contribution of Verbal Comprehension subtests to the Full Scale IQ is reduced from 50% to 40%, Visual Spatial from 50% to 40%, and Working Memory is changed from 0% to 20%. While an additional subtest contributes to the Full Scale IQ, the testing time necessary to obtain it is unchanged.

For ages 4:0 to 7:7, the changes relative to WPPSI-III are more extensive. The Full Scale IQ is derived using one fewer subtest (i.e., six) relative to the WPPSI-III, and the testing time to obtain the Full Scale IQ is shorter. The contribution of Verbal Comprehension subtests to the Full Scale IQ is decreased from 43% to 33%, and the contribution of Visual Spatial subtests to the Full Scale IQ is slightly higher (17% versus 14%), although Block Design remains the only contributing subtest from that domain. The contribution of Fluid Reasoning is also reduced (from 29% to 17%), as Picture Concepts is no longer core to the Full Scale IQ. As with younger children, one Working Memory subtest contributes to the Full Scale IQ, although the relative contribution of Working Memory is 17%. One Processing Speed subtest continues to contribute to the Full Scale IQ; however, the relative contribution of Processing Speed is slightly higher (17% versus 14%) because there are fewer subtests contributing to the Full Scale IQ. Three subtests (50% of the total) are shared across the WPPSI-III and WPPSI-IV Full Scale IQs due to different selections for core subtests (e.g., Similarities instead of Word Reasoning and Vocabulary) and replacements for dropped subtests (e.g., Bug Search instead of Coding).

There are other obvious changes to the test content and structure. New constructs are measured with the new Working Memory Index and subtests and the new Visual Spatial Index for ages 4:0 to 7:7, and the new Fluid Reasoning index for ages 4:0 to 7:7. The Verbal Comprehension subtests now all have initial picture items to ensure children with expressive issues and shy children who initially are hesitant to respond verbally can experience some success, and the new Processing Speed subtests are more developmentally appropriate.

The relation of the WPPSI-IV to the WPPSI-III was examined in 246 children aged 2:6 to 7:3 (mean age of 4.8). The tests were administered in counterbalanced order with a mean testing interval of 22 days and a range of 13 to 54 days (Wechsler, 2012). The sample contained representation from a variety of children from different races/ethnicities, parent education levels, and U.S. geographic regions, and roughly comprised half female and half male. Table 1.1 presents the mean composite score on each version, the standard difference, and the corrected correlation coefficients.

The overall correlation indicates that the Full Scale IQs for the two versions are the most highly correlated of all composite scores (.86), followed by the General Language Composite–Verbal Acquisition Index (.85), the Verbal Comprehension Index (.84), the Performance IQ–Fluid Reasoning Index (.76), and the Performance IQ–Visual Spatial Index (.71).

The Processing Speed Quotient–Processing Speed Index correlation was the lowest among the composite scores (.65), which is not an unexpected result due to the complete replacement of subtests across the two versions. Regardless of the extensive revisions to the test, the high correlation coefficient of .86 indicates the Full Scale IQ continues to measure the same construct.

As presented in Table 1.1, the average WPPSI-IV Full Scale IQ is 3.3 points lower than the WPPSI-III Full Scale IQ. The Verbal IQ–Verbal Comprehension

Table 1.1 Comparison of WPPSI-III and WPPSI-IV Scores

	WPPSI-III		WPPSI-IV			
Score	Mean	SD	Mean	SD	Standard Difference	Corrected Correlation
VCI–VIQ	103.4	13.5	100.9	12.8	.19	.84
VSI–PIQ	104.9	13.7	102.6	13.3	.17	.71
FRI–PIQ	105.4	13.6	102.1	12.5	.25	.76
WMI			100.4	12.9		
PSI–PSQ	107.0	12.0	101.1	12.8	.48	.65
Full Scale IQ	105.0	13.5	101.7	13.0	.25	.86
VAI–GLC	104.6	13.4	101.7	12.8	.22	.85
NVI			102.1	12.9		
GAI			101.9	13.7		
CPI			100.4	11.7		

Source: Adapted from Table 5.5 of the *WPPSI-IV Technical and Interpretive Manual*.

Index, Performance IQ–Visual Spatial Index, and Performance IQ–Fluid Reasoning Index differences are similar to those observed for the Full Scale IQs, although the correlations are somewhat lower than those of the Full Scale IQ or the verbal composites. This likely occurs because the Performance IQ was split into two separate composites, so less construct overlap occurs across the two tests for these scores. These differences are in the direction and of the same size as expected according to the Flynn effect (Flynn, 2007), which predicts that observed scores on an older test become higher over time due to outdated norms. Hence, children score somewhat lower on the WPPSI-IV relative to the WPPSI-III, but the scores more accurately reflect the child's intellectual functioning because the norms are based on samples from the current population.

The largest standard difference was observed for the Processing Speed Quotient–Processing Speed Index (.48, with a difference of almost 6 points). This is not surprising, due to the new subtests that contribute to the WPPSI-IV Processing Speed Index and the relatively poorer floors of the WPPSI–III Processing Speed subtests. Children who could perform the tasks rather easily obtained higher scaled scores than on the corresponding WPPSI-IV subtests.

STANDARDIZATION AND PSYCHOMETRIC PROPERTIES

The WPPSI-IV normative information is based on a national sample of 1,700 children. It was collected from December 2010 through May 2012. Children were selected to match census proportions from 2010 U.S. Census data and the sample is stratified according to age, sex, race/ethnicity, parent education level, and U.S. geographic region. Nine age groups were created, with 200 children in each of eight age groups from 2:6 to 6:11 and 100 children in the 7:0 to 7:7 age group.

Reliability

The average reliability coefficient for the Full Scale IQ across the nine age groups was excellent, at .96 overall, with a range of .95 to .96 across the age groups. The primary index scores have overall reliability coefficients ranging from .86 for the Processing Speed Index to .94 for the Verbal Comprehension Index. The reliability coefficients ranged from .85 to .95 for the primary index scores at the individual age-group level. The subtest reliability coefficients ranged from .75 for Animal Coding to .93 for Similarities. At the age-group level, the subtest reliability coefficients ranged from .71 (for Animal Coding at the youngest ages) to .95 (for Similarities at the youngest age).

A subset of the normative sample ($N = 172$) provided retest reliability data, with an average of 23 days between the first and second testing. Results showed similar stability coefficients across the three age ranges in the study (2:6 to 3:11; 4:0 to 5:5, and 5:6 to 7:7). The average stability coefficients across all ages for composite scores ranged from .86 for the Processing Speed Index to .93 for the Full Scale IQ and the General Ability Index. The highest overall average subtest stability coefficient was .87 for Similarities, and the lowest was .75 for Zoo Locations and Animal Coding.

The average reliability coefficients for subtest, process, and composite scores, by age band and for all applicable ages, appear in Rapid Reference 1.11. Both internal consistency and test-retest stability coefficients are presented.

Loadings on the General Factor

General intelligence, or *g* (Spearman, 1927), can be derived by several methods. For the purposes of this book, *g* is calculated using the subtest factor loadings on the first unrotated factor in a principal components analysis. Factor loadings of .70 or greater are classified as good measures of *g*, loadings of .50 to .69 are classified as fair, and loadings below .50 are classified as poor. Squaring the subtest *g* loading provides the proportion of variance attributable to *g*.

Rapid Reference 1.11

Average Reliability Coefficients of Subtest, Process, and Composite Scores

Subtest/ Composite Score	2:6–3:11 Internal Consistency	4:0–7:7 Internal Consistency	All Applicable Ages Internal Consistency	All Applicable Ages Test-Retest Stability
IN	.91	.88	.89	.83
SI	—	.93	.93	.87
VC	—	.89	.89	.84
CO	—	.91	.91	.84
RV	.91	.90	.90	.79
PN	.89	.88	.88	.83
BD	.85	.85	.85	.81
OA	.85	.85	.85	.77
MR	—	.90	.90	.82
PC	—	.89	.89	.79
PM	.91	.90	.91	.80
ZL	.90	.84	.86	.75
BS	—	—	—	.83
CA	—	—	—	.76
AC	—	—	—	.75
VCI	.94	.94	.94	.89
VSI	.89	.90	.89	.86
FRI	—	.93	.93	.88
WMI	.93	.91	.91	.87
PSI	—	—	—	.86
FSIQ	.96	.96	.96	.93
VAI	.94	.93	.93	.86
NVI	.94	.95	.95	.90
GAI	.95	.95	.95	.93
CPI	—	.92	.92	.89

Note. Abbreviations are IN = Information, SI = Similarities, VC = Vocabulary, CO = Comprehension, RV = Receptive Vocabulary, PN = Picture Naming, BD = Block Design, OA = Object Assembly, MR = Matrix Reasoning, PC = Picture Concepts, PM = Picture Memory, ZL = Zoo Locations, BS = Bug Search, CA = Cancellation, AC = Animal Coding, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, VAI = Vocabulary Acquisition Index, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.
Source: Data are from the *Technical and Interpretive Manual* Tables 4.1 and 4.5.

Lichtenberger and Kaufman (2004) noted that the meaning of *g* loadings and of the concept of general intelligence has been the subject of much discussion and debate. That debate continues to the present day (S. B. Kaufman et al., 2012; Reynolds, 2013; te Nijenhius, van Vianen, & van der Flier, 2007). It therefore is important, as Lichtenberger and Kaufman state, not to interpret a subtest with a good *g* loading in isolation as representative of the child's general intellectual ability.

The subtest *g* loadings, strength of each subtest as a measure of *g*, and proportion of variance for each subtest attributed to *g*, by age band, is provided in Rapid Reference 1.12. The subtests are listed in descending order with respect to *g* loading.

Rapid Reference 1.12

Subtest *g* Loadings, Strength as Measures of *g*, and Proportions of Variance Attributed to *g*, by Age Band

Subtest	g Loading	Strength as Measure of g	Proportion of Variance Attributed to g
		Ages 2:6–3:11	
IN	.80	good	.65
PN	.78	good	.62
RV	.77	good	.59
PM	.69	fair	.48
BD	.67	fair	.45
OA	.63	fair	.40
ZL	.55	fair	.30
		Ages 4:0–7:7	
SI	.78	good	.60
IN	.77	good	.60
VC	.76	good	.58
CO	.73	good	.53
PN	.73	good	.53
RV	.71	good	.51
MR	.68	fair	.46
BD	.68	fair	.46
OA	.64	fair	.42

(continued)

Subtest	g Loading	Strength as Measure of g	Proportion of Variance Attributed to g
		Ages 4:0–7:7	
PM	.63	fair	.40
BS	.62	fair	.38
PC	.61	fair	.37
AC	.56	fair	.31
ZL	.55	fair	.30
CA	.51	fair	.26

Note. Abbreviations are IN = Information, SI = Similarities, VC = Vocabulary, CO = Comprehension, RV = Receptive Vocabulary, PN = Picture Naming, BD = Block Design, OA = Object Assembly, MR = Matrix Reasoning, PM = Picture Memory, ZL = Zoo Locations, BS = Bug Search, CA = Cancellation, AC = Animal Coding.

All *g* loadings of .70 or above are considered good, .50 to .69 are considered fair, and loadings below .50 are considered poor.

None of the WPPSI-IV subtests are poor measures of *g*; all are good or fair. For younger and older children, the strongest *g* loadings occur on the Verbal Comprehension subtests. All Verbal Comprehension subtests are good measures of *g*, and all other subtests are fair measures of *g*. This pattern is similar to that observed for the WPPSI-III (Lichtenberger & Kaufman, 2004).

For children aged 2:6 to 3:11, Information has the highest *g* loading (.80), followed by Picture Naming (.78) and Receptive Vocabulary (.77). For children aged 4:0 to 7:7, the core subtests for the Verbal Comprehension Index have the highest *g* loadings; Similarities is the highest (.78) by a slight margin (Information is .77). For both Information and Similarities, 60% of the variance is attributed to *g*. The next highest *g* loadings occur on the subtests that require verbal expression in at least some responses: Vocabulary (.76) and Comprehension (.73), and Picture Naming (.73). Among the Verbal Comprehension subtests, Receptive Vocabulary has the lowest *g* loading for this age band (.71).

For ages 2:6 to 3:11, the other three subtests that are core to the Full Scale IQ have the next-highest *g* loadings. Of these, Picture Memory has the highest *g* loading (.69), followed by Block Design (.67) and Object Assembly (.63). The lowest *g* loading occurs for Zoo Locations (.55), which attributes 30% of its variance to *g*.

For ages 4:0 to 7:7, the remaining subtests that are core to the Full Scale IQ, Matrix Reasoning, (.68), Block Design (.68), Picture Memory (.63), and Bug

Search (.62) occupy four of the subsequent five positions in descending order of *g* loading. Only Object Assembly (.64), which is core to the Visual Spatial Index, is present among those four core Full Scale IQ subtests. Picture Concepts (.61) has the next highest *g* loading, followed by two of the Processing Speed subtests, Animal Coding (.56) and Cancellation (.51), and Zoo Locations (.55). Zoo Locations produces the same relatively low *g* loading in both age bands. Of all WPPSI-IV subtests, Cancellation has the lowest *g* loading. However, its *g* loading is superior to other versions of Cancellation from WAIS-IV (.38; Lichtenberger & Kaufman, 2013) and WISC-IV (.25; Flanagan & Kaufman, 2009). From the beginning of WPPSI-IV development, we hypothesized that the Cancellation *g* loading would be higher for very young children relative to school-age children because classification of simple objects is a more challenging cognitive task at younger ages, and requires cognitive flexibility that not all young children have attained.

Subtest Specificities

The unique proportion of reliable variance of each subtest, or the proportion of subtest variance unrelated to measurement error and specific to that subtest (i.e., not shared with other subtests), is termed the *subtest specificity*. While interpretation at the subtest level is not recommended, the specificities are nevertheless useful when attempting to understand subtest performance. Lichtenberger and Kaufman (2004) suggest that if about 25% of a subtest's variance is specific, and the specific variance exceeds the subtest's error variance, the specificity associated with that subtest is sufficiently meaningful (ample). Lichtenberger and Kaufman's interpretive approach at that time involved grouping shared abilities, and only interpreting unique subtest abilities as a last resort. Their approach evolved subsequently, to focus much more on interpreting index-level strengths and weaknesses (Lichtenberger & Kaufman, 2013). Their more current work does not emphasize specificities for that reason.

We include the subtest specificities because we view them as another piece of information that facilitates a more complete understanding of the WPPSI-IV subtests. Evaluating their unique contribution to the battery and to the composite scores from which they are derived and understanding the relations of specific variance and error variance only adds richness and expertise to interpretation. Conceptually, we view the subtest specificities as an important aspect of battery selection: We believe specificity is an indicator of the lack of redundancy across selected measures. Flanagan and colleagues' cross-battery approach (Flanagan, Ortiz, & Alfonso, 2013) also conceptualizes lack of redundancy as important in evaluations. For example, their approach classifies measurement of a CHC broad ability as inadequate if at least two CHC narrow abilities are not represented.

To obtain the subtest specificity, the squared multiple correlation (from maximum-likelihood factor analysis with varimax rotation) is subtracted from the reliability for that subtest. To obtain the subtest's error variance, subtest's reliability is subtracted from 1.The subtest specificities for each age band, along with the error variance and strength of subtest specificity for each, are provided in Rapid Reference 1.13. The subtests are presented in descending order according to specificity value.

As seen in Rapid Reference 1.13, all subtests have ample specificity for both age bands. For children aged 2:6 to 3:11, the highest specificity value occurs for Zoo Locations, which does not contribute to Full Scale IQ. The highest subtest

Rapid Reference 1.13

Subtest Specificities, by Age Band

Subtest	Squared Multiple Correlation*	Specificity	Error Variance	Strength of Subtest Specificity
Ages 2:6–3:11				
ZL	.19	.71	.10	ample
OA	.26	.59	.15	ample
PM	.34	.57	.09	ample
BD	.30	.55	.15	ample
RV	.44	.47	.09	ample
IN	.54	.37	.09	ample
PN	.52	.37	.11	ample
Ages 4:0–7:7				
PC	.32	.57	.11	ample
PM	.35	.55	.10	ample
ZL	.30	.54	.16	ample
MR	.41	.49	.10	ample
OA	.38	.47	.15	ample
CA	.29	.47	.24	ample
RV	.47	.43	.10	ample
BD	.42	.43	.15	ample
BS	.43	.40	.17	ample
AC	.37	.38	.25	ample

Subtest	Squared Multiple Correlation*	Specificity	Error Variance	Strength of Subtest Specificity
		Ages 4:0–7:7		
CO	.54	.37	.09	ample
PN	.54	.34	.12	ample
SI	.60	.33	.07	ample
VC	.58	.31	.11	ample
IN	.58	.30	.12	ample

* The squared multiple correlations are from maximum-likelihood factor analysis with varimax rotation.

Note. Abbreviations are IN = Information, SI = Similarities, VC = Vocabulary, CO = Comprehension, RV = Receptive Vocabulary, PN = Picture Naming, BD = Block Design, OA = Object Assembly, MR = Matrix Reasoning, PM = Picture Memory, ZL = Zoo Locations, BS = Bug Search, CA = Cancellation, AC = Animal Coding.

Ample specificity = specific variance that is at least 25% of total subtest variance, and greater than the subtest's error variance. Adequate specificity = specific variance that is 15%–24% of total subtest variance, and greater than the subtest's error variance. Inadequate specificity = specific variance less than 15% of total variance, or subtest's error variance exceeds specific variance.

To obtain the subtest specificity, the squared multiple correlation (from maximum-likelihood factor analysis with varimax rotation) is subtracted from the reliability for that subtest. To obtain the subtest's error variance, subtract the subtest's reliability from 1.

specificity for ages 4:0 to 7:7 is observed for Picture Concepts, followed by Picture Memory and Zoo Locations. In general, more highly *g*-loaded subtests tend to have lower subtest specificity values, and those with lower *g* loadings tend toward higher subtest specificity values. The Verbal Comprehension subtests, with the exception of Receptive Vocabulary in the older age band, tend to produce lower subtest specificities relative to other subtests. Subtests that rely on visual stimuli tend to show higher subtest specificities.

COMPREHENSIVE TEST REFERENCES

The *WPPSI-IV Administration and Scoring Manual* (Wechsler, 2012) and the *WPPSI-IV Technical and Interpretive Manual* (Wechsler, 2012) currently provide the most detailed information about the WPPSI-IV. These manuals review the scale's development, subtest descriptions, item- and subtest-level administration and scoring rules, standardization, and evidence of reliability and validity. Rapid Reference 1.14 provides basic information on the WPPSI-IV and the test publisher, Pearson.

Essentials of WPPSI-III Assessment (Lichtenberger & Kaufman, 2004) provides complete information about administration, scoring, and interpretation of the

Rapid Reference 1.14

- Wechsler Preschool and Primary Scale of Intelligence–Fourth Edition (WPPSI-IV)
- Author: David Wechsler
- Publication Date: 2012
- Age Range: 2:6 to 7:7
- What the Test Measures: verbal comprehension, visual spatial processing, fluid reasoning (for ages 4:0 to 7:7 only), working memory, processing speed (for ages 4:0 to 7:7 only), and general intellectual ability
- Administration Time: Full Scale IQ—24 to 29 minutes for younger children, 31 to 32 minutes for older children; Primary Index Subtests—29 to 35 minutes for younger children, 58 to 62 minutes for older children
- Qualification of Examiners: C level
- Publisher: Pearson
- 5601 Green Valley Drive
- Bloomington, MN 55437
- Customer Service: (800) 627-7271
- http://www.PsychCorp.com
- Product Number: 0158984900
- Price: WPPSI-IV Basic Kit: Includes (list materials) in a box. $1,120 (in box); $1,200 (in a rolling bag), $1,190 (in a hard case).

prior edition, the WPPSI-III (Wechsler, 2003). *Assessment of Children: Cognitive Foundations*, fifth edition, and its resource guide that accompanies the book (Sattler, 2008) provides an in-depth review of WPPSI-III administration and scoring, as well as relevant research on prior editions.

TEST YOURSELF

1. **Which of the following WPPSI-IV subtests is a *core* subtest that is used to compute FSIQ for ages 2:6 to 7:7?**
   (a) Cancellation
   (b) Picture Naming
   (c) Zoo Locations
   (d) Picture Concepts
   (e) Picture Memory

2. **What major structural change was implemented from the WPPSI-III to the WPPSI-IV?**
3. **For children aged 4:0 to 7:7, which two primary index scores replace the Performance IQ from WPPSI-III?**
4. **Which is a retained subtest from WPPSI-III?**
   (a) Picture Concepts
   (b) Picture Memory
   (c) Cancellation
   (d) Bug Search
   (e) Zoo Locations
5. **Which subtest is *not* a measure of Processing Speed?**
   (a) Cancellation
   (b) Animal Coding
   (c) Zoo Locations
   (d) Bug Search
6. **Which is *not* an Ancillary Index Score?**
   (a) General Ability Index
   (b) Vocabulary Acquisition Index
   (c) Cognitive Proficiency Index
   (d) Fluid Reasoning Index
   (e) Nonverbal Index
7. **Which is *not* a primary Index Score?**
   (a) Working Memory Index
   (b) Nonverbal Index
   (c) Visual Spatial Index
   (d) Fluid Reasoning Index
   (e) Verbal Comprehension Index
8. **Which primary index score includes the subtests that have the lowest subtest reliability coefficients for children ages 4:0 to 7:7?**
   (a) Visual Spatial Index
   (b) Working Memory Index
   (c) Processing Speed Index
   (d) Verbal Comprehension Index
   (e) Fluid Reasoning Index
9. **In general, more highly *g*-loaded subtests tend to have higher subtest specificity values, and those with lower *g* loadings tend toward lower subtest specificity values.**

   True or False?

(*continued*)

10. Which primary index contains the subtests that have the highest overall g loadings?

Answers: 1. e; 2. Verbal IQ and Performance IQ are no longer in use, and PIQ has been replaced by the Visual Spatial Index and the Fluid Reasoning Index; 3. the Visual Spatial Index and the Fluid Reasoning Index; 4. a; 5. c; 6. d; 7. b; 8. c; 9. False; 10. Verbal Comprehension Index

REFERENCES

Albert, M. L. (1973). A simple test of visual neglect. *Neurology*, *23*, 658–664.

Ammons, R. B., & Ammons, H. S. (1948). *The Full Range Vocabulary Test*. New Orleans, LA: Authors.

Binet, A. (1911). Nouvelles recherches sur la mesure du niveau intellectuel chez les enfants d'école. *L'Année Psychologique*, *14*, 145–201.

Binet, A., & Simon, T. (1905). Méthods nouvelles pour le diagnostic du niveau intellectual des anormaux. *L'Année Psychologique*, *11*, 191–244.

Binet, A., & Simon, T. (1908). Le développement de l'intelligence chez les enfants. *L'Année Psychologique*, *14*, 1–90.

Boake, C. (2002). From the Binet-Simon to the Wechsler-Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology*, *24*, 383–405.

Carlson, L., & Reynolds, C. R. (1981). Factor structure and specific variance of the WPPSI subtests at six age levels. *Psychology in the Schools*, *18*(1), 48–54.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.

Carroll, J. B. (2012). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 883–890). New York, NY: Guilford Press.

Cattell, P. (1940). *The measurement of intelligence of infants and young children*. New York, NY: Psychological Corporation.

Coalson, D. L., Raiford, S. E., Saklofske, D. H., & Weiss, L. G. (2010). In L. G. Weiss, D. H. Saklofske, D. L. Coalson, & S. E. Raiford (Eds.), *WAIS-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 3–23). Amsterdam, The Netherlands: Elsevier Academic Press.

Elliott, C. (2007). *Differential ability scales* (2nd ed.). San Antonio, TX: Pearson.

Flanagan, D. P., Alfonso, V. C., & Ortiz, S. O. (2012). The cross-battery assessment approach. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 459–483). New York, NY: Guilford Press.

Flanagan, D. P., Alfonso, V. C., Ortiz, S. O., & Dynda, A. M. (2010). Integrating cognitive assessment in school neuropsychological evaluations. In D. C. Miller (Ed.), *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention* (pp. 101–140). Hoboken, NJ: Wiley.

Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC-IV assessment* (2nd ed.). Hoboken, NJ: Wiley.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3rd ed.). Hoboken, NJ: Wiley.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge, England: Cambridge University Press.

Ford, L., Kozey, M. L., & Negreiros, J. (2012). Cognitive assessment in early childhood: Theoretical and practical perspectives. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 585–622). New York, NY: Guilford Press.

Gesell, A., & Amatruda, C. S. (1941). *Developmental diagnosis: Normal and abnormal child development, clinical methods, and pediatric applications*. New York, NY: Hoeber.

Goddard, H. H. (1916). *The development of intelligence in children (the Binet-Simon Scale)*. Baltimore, MD: Williams & Wilkins.

Goldstein, G., & Beers, S. R. (Eds.). (2003). *Comprehensive handbook of psychological assessment: Vol. 1. Intellectual and neuropsychological assessment*. Hoboken, NJ: Wiley.

Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4th ed.). Hoboken, NJ: Wiley.

Gyurke, J. S., Stone, B. J., & Beyer, M. (1990). A confirmatory factor analysis of the WPPSI–R. *Journal of Psychoeducational Assessment*, *8*, 15–21.

Horn, J. L., & Blankson, A. N. (2012). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 73–98). New York, NY: Guilford Press.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53–91). New York, NY: Guilford Press.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman assessment battery for children* (2nd ed.). Bloomington, MN: Pearson.

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive *g* and academic achievement *g* one and the same *g*? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence*, *40*, 123–138. doi: 10.1016/j.intell.2012.01.009

Kelley M. F., & Surbeck, E. (2007). History of preschool assessment. In B. A. Bracken & R. J. Nagle (Eds.), *Psychoeducational assessment of preschool children* (pp. 3–28). Mahwah, NJ: Erlbaum.

Kirkpatrick, E. A. (1909). Studies in development and learning. *Archives of Psychology*, *12*, 1–101.

Kohs, S. C. (1923). *Intelligence measurement: A psychological and statistical study based upon the block design tests*. New York, NY: Macmillan.

Kuhlmann, F. (1912). A revision of the Binet-Simon system for measuring the intelligence of children. *Journal of Psycho-Asthenics Monographs Supplement*, *1*(1), 1–41.

Kuwajima, M., & Sawaguchi, T. (2010). Similar prefrontal cortical activities between general fluid intelligence and visuospatial working memory tasks in preschool children as revealed by optical topography. *Experimental Brain Research*, *206*, 381–397.

Leiter, R.G. (1948). *International Performance Scale*. Chicago, IL: Stoelting.

Lichtenberger, E. O., & Kaufman, A. S. (2004). *Essentials of WPPSI®–III assessment*. Hoboken, NJ: Wiley.

Lichtenberger, E. O., & Kaufman, A. S. (2013). *Essentials of WAIS®–IV assessment* (2nd ed.). Hoboken, NJ: Wiley.

Locke, J. (1692). Some thoughts concerning education. In C. W. Eliot (Ed.), *Harvard Classics*, *37*(1), New York, NY: Collier, 1909–1914; Bartleby.com, 2001. www.bartleby.com/37/1/

McCarthy, D. (1972). *The McCarthy Scales of Children's Abilities*. San Antonio, TX: Psychological Corporation.

Miller, D. C. (2010). School neuropsychological assessment and intervention. In D. C. Miller (Ed.), *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention* (pp. 81–100). Hoboken, NJ: Wiley.

Miller, D. C. (2013). *Essentials of school neuropsychological assessment*. Hoboken, NJ: Wiley.

Morgan, K. E., Rothlisberg, B. A., McIntosh, D. E., & Hunt, M. S. (2009). Confirmatory factor analysis of the KABC–II in preschool children. *Psychology in the Schools*, *46*(6), 515–525.

Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: International University Press.

Pintner R. & Paterson, D. G. (1917). *A scale of performance tests*. New York, NY: Appleton.

Raven, J. C. (1941). Standardization of progressive matrices, 1938. *British Journal of Medical Psychology*, *19*, 137–150.

Reynolds, M. R. (2013). Interpreting the *g* loadings of intelligence test composite scores in light of Spearman's law of diminishing returns. *School Psychology Quarterly*, *28*, 63–76.

Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds intellectual assessment scales*. Lutz, FL: Psychological Assessment Resources.

Roid, G. H. (2005). *Early Childhood Stanford-Binet intelligence scales* (5th ed). Austin, TX: Pro-Ed.

Rousseau, J. (1762 /1979). *Emile, or On education* (A. Bloom, Trans.). New York, NY: Basic Books.

Sattler, J. M. (2008). *Assessment of children: Cognitive foundations* (5th ed.). San Diego, CA: Author.

Schneider, B. H., & Gervais, M. D. (1991). Identifying gifted kindergarten students with brief screening measures and the WPPSI-R. *Journal of Psychoeducational Assessment*, *9*, 201–208.

Schneider, W., Schumann-Hengsteler, R., & Sodian, B. (Eds.). (2005). *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind*. Mahwah, NJ: Erlbaum.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99–144). New York, NY: Guilford Press.

Silverstein, A. B. (1969). An alternative factor analytic solution for Wechsler's intelligence scales. *Educational and Psychological Measurement*, *29*, 763–767.

Spearman, C. (1927). *The abilities of man*. London, England: MacMillan.

Sternberg, R. J. (Ed.). (2000). *Handbook of intelligence*. New York, NY: Cambridge University Press.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654.

Stone, B. J., Gridley, B. E., & Gyurke, J. S. (1991). Confirmatory factor analysis of the WPPSI–R at the extreme of the age range. *Journal of Psychoeducational Assessment*, *9*, 263–270.

te Nijenhius, J., van Vianen, A. E. M., & van der Flier, H. (2007). Score gains on *g*-loaded tests: No *g*. *Intelligence*, *35*, 283–300.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale*. Boston, MA: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence*. Boston, MA: Houghton Mifflin.

Ward, K. E., Rothlisberg, B. A., McIntosh, D. E., & Bradley, M. H. (2011). Assessing the SB–V factor structure in a sample of preschool children. *Psychology in the Schools*, *48*(5), 454–463.

Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3–55). New York, NY: Guilford Press.

Wechsler, D. (1939). *Wechsler-Bellevue intelligence scale*. New York, NY: Psychological Corporation.

Wechsler, D. (1944). *The measurement of adult intelligence* (3rd ed.). Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1949). *Wechsler intelligence scale for children*. New York, NY: Psychological Corporation.

Wechsler, D. (1950). Cognitive, conative, and non-intellective intelligence. *American Psychologist*, *5*(3), 78–83.

Wechsler, D. (1955). *Wechsler adult intelligence scale*. New York, NY: Psychological Corporation.

Wechsler, D. (1967). *Wechsler preschool and primary scale of intelligence*. New York, NY: Pearson.

Wechsler, D. (1989). *Wechsler preschool and primary scale of intelligence* (revised). San Antonio, TX: Psychological Corporation.

Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence* (3rd ed.). San Antonio, TX: Pearson.

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio, TX: Pearson.

Wechsler, D. (2008). *Wechsler adult intelligence scale* (4th ed.). San Antonio, TX: Pearson.

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence* (4th ed.). Bloomington, MN: Pearson.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *Woodcock-Johnson tests of cognitive ability* (3rd ed.) *normative update*. Itasca, IL: Riverside.

Wortham, S. C. (2012). *Assessment in early childhood education* (6th ed.). New York, NY: Pearson.