Metadata of the chapter that will be visualized online

Chapter Title	100 Years of Intelligence Testing: Moving from Traditional IQ to Second-Generation Intelligence Tests	
Copyright Year	2015	
Copyright Holder	Springer Science+Business Media New York	
Corresponding Author	Family Name	**Naglieri**
	Particle	
	Given Name	**Jack A.**
	Suffix	
	Address	6622 White Post Road, Centreville, VA, 20121, USA
	Email	jnaglieri@gmail.com
Abstract	This chapter provides an historical perspective on intelligence and IQ testing from 1917 to present day. Traditional IQ tests are compared to second generation measures on several important dimensions including profiles of test scores, fairness to minority groups, and relevance to academic interventions. The PASS neurocognitive theory of intelligence as measured by the Cognitive Assessment System (First and Second Editions) is described as having the strongest research base and the most viable approach.	
Keywords (separated by "-")	Intelligence - IQ - Wechsler - Planning - Attention - Simultaneous - Successive (PASS) theory of intelligence - Race and ethnic differences	

100 Years of Intelligence Testing: Moving from Traditional IQ to Second-Generation Intelligence Tests

20

Jack A. Naglieri

"Do not go where the path may lead, go instead where there is no path and leave a trail."

-Ralph Waldo Emerson

Context

April 6, 1917, is remembered as the day the United States entered World War I. On that same day a group of psychologists held a meeting in Harvard University's Emerson Hall to discuss the possible role they could play with the war effort [AU1] (Yerkes 1921). The group agreed that psychological knowledge and methods could be of importance to the military and utilized to increase the efficiency of the Army and Navy personnel. The group included Robert Yerkes, who was also the president of the American Psychological Association. Yerkes made an appeal to members of APA who responded by providing a group of psychologists to assist with the war effort. Members from APA were joined by psychologists of the National Research Council, the National Academy of Sciences, and the American Association for the Advancement of Science; and a number of committees were organized to develop effective measures of ability.

One group of psychologists whose task was to begin identifying possible tests met at the Training School in Vineland, New Jersey, on May 28. The committee considered many types of group tests and several that Arthur S. Otis developed when working on his doctorate under Lewis Terman at Stanford University. The goal was to find tests that could efficiently evaluate a wide variety of men, be easy to administer in the group format, and be easy to score. By June 9, 1917, the materials were ready for an initial trial. Men who had some educational background and could speak English were administered the verbal and quantitative (Alpha) tests and those that could not read the newspaper or speak English were given the Beta tests (today described as nonverbal).

The Alpha tests were designed to measure general information (e.g., how many months are there in a year?), common sense (e.g., why do we use stoves?), and verbal knowledge (synonym/antonyms, verbal analogies, number series, disarranged sentences) (e.g., determine if a group of words could be sequenced to make a true statement) and to determine how well the examinee could follow verbal directions (e.g., draw a line from circle 3 to circle 6). The Beta tests included completion of a maze, construction of a design using blocks, number symbol association, identifying what is missing in a picture, and copying geometric shapes. Why two tests? Because, as Yoakum and Yerkes (1920) clearly stated, the Alpha test was an appropriate measure of intelligence for men who could read and write English sufficiently. The Beta tests were intended for those who had difficulty reading or spoke English

J.A. Naglieri, Ph.D. (✉)
[AU2] ~~6622 White Post Road, Centreville, VA 20121, USA~~
e-mail: jnaglieri@gmail.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_20,

poorly as well as those who were illiterate or not able to understand English (p. 51). The testing procedures ensured that men "who fail in alpha are sent to beta in order that injustice by reason of relative unfamiliarity with English may be avoided (Yoakum & Yerkes, p. 19)."

By July 7, 1917, the initial development of tests to measure intellectual ability and a study involving 400 cases was completed. The data obtained from testing sites in Indiana, Tennessee, and Syracuse and Brooklyn, New York, was shipped to the statistical unit in New York City for data analysis. Statistical analysis was directed by E. L. Thorndike with assistants A. S. Otis and L. L. Thurstone. The report of their analysis completed on July 20, 1917, showed that the tests could be appropriate to (a) "aid in segregating and eliminating the mentally incompetent, (b) classify men according to their mental ability; and (c) assist in selecting competent men for responsible positions" (Yerkes 1921, p. 19). Thus, July 20, 1917, could be considered the birth date of the verbal, quantitative, and nonverbal IQ test format which will be called traditional IQ in the remainder of this chapter.

By early 1918 a group of about 260 men trained in the Medical Corps School for Military Psychology began using the Army Alpha and Beta tests. Among them was the 22-year-old David Wechsler (1896–1981) who arrived at Fort Logan, Texas, in August that year. Wechsler, like Yerkes, who wrote in the Introduction to the *Psychological Examining in the United States Army* (1921), noticed "the educational, industrial, and significance of the methods [Alpha and Beta] (p. 5)". Wechsler's adaptation some 20 years later resulted in the Wechsler-Bellevue Intelligence Scale (1939) which also contained verbal and quantitative (the Alpha) and performance (the Beta) test questions. These tests would have been familiar to the founders of the company that ultimately published Wechsler's test (the Psychological Corporation), because many years earlier (in 1917), R. S. Woodworth was the chief examiner at the Brooklyn testing site collecting initial data on the Army Alpha and Beta; Thorndike was responsible for the analysis of those data and R. Cattell was initially on the original National Research Council meeting that led to the development of the measures.

The transition from Army Alpha and Beta to Wechsler IQ as we know it today is the result of the work of those psychologists who first met on April 6, 1917, and those that developed, validated, and used the Army Alpha and Beta tests. It would likely have been well beyond the expectations of Major Lewis M. Terman, Captain Edwin G. Boring, and related personnel including R. S. Woodworth, E. L. Thorndike, A. S. Otis, and L. L Thurstone that the work they did would become the most widely used approach to measuring intelligence in history. And in addition, their work (described in the book *Army Mental Tests* by Yoakum and Yerkes 1920) would define the structure of individual-(e.g., Wechsler Scales) and group-administered (Otis-Lennon 1979) IQ tests for the next 100 years.

The evolution of traditional IQ tests from their birth in 1917 has been defined by the many revisions of Wechsler and Otis-Lennon tests, the most current version of the latter being number 8 and the forthcoming fifth edition of the Wechsler Intelligence Scale for Children (WISC-V, Wechsler 2014). Ironically, there have been many revisions of these tests long after the authors died (Otis in 1964 and Wechsler in 1981). What has evolved is a body of knowledge about these tests and how to interpret them. For example, the first book on the development and interpretation of the Wechsler Scales was the 1939 publication of *Wechsler's Measurement and Appraisal of Adult Intelligence*. The fifth edition of that book was published (authored by Matarazzo) in 1976. Shortly thereafter, Alan Kaufman's 1979 book, *Intelligence Testing with the WISC-R,* provided practitioners with a wealth of information about both psychometrically sound and clinically relevant interpretation methods. This was followed by *Intelligence Testing with the WISC-III* (Kaufman 1994) and *WISC-IV* (Flanagan and Kaufman 2004). Other authors have also provided books on how to extract information about Wechsler's scales (e.g., Prifitera and Saklofske 1998; Weiss et al. 2006). All this effort has been focused on ways to interpret a test based on measures first assembled in 1917.

Despite the widespread use and acceptance of the traditional verbal/quantitative (Alpha) and nonverbal (Beta) IQ approach to intelligence, it is important to ask if this approach meets the demands of psychology and education today. Are the tests appropriate for diverse populations? Does the information traditional IQ tests yield assist in planning instruction and academic interventions? How well do these tests evaluate the intellectual component of a specific learning disability? Perhaps most importantly, should we continue to use IQ tests which were devoid of a theoretical foundation from the time they were first introduced more than 100 years ago (Naglieri and Kaufman 2008)? The purpose of this chapter is to provide answers to these questions and evidence for an alternative to traditional IQ based on brain function. The short answer is yes; our understanding of intelligence has evolved and better ways to measure it are now available.

Starting Over

The evolution of intelligence tests was stimulated by the publication of two *second-generation* ability tests. First was the publication of the *Kaufman Assessment Battery for Children* (Kaufman and Kaufman 1983). This test was based on a blend of perspectives about what intelligence may be. What the Kaufman's did that was most important was to tie a test of intelligence with a two-dimensional model of ability conceptualized within a cognitive processing context. Another very important aspect of the KABC and the second edition (K-ABC-II, Kaufman and Kaufman 2004) was the shift away from organizing their test based on the verbal, quantitative, and nonverbal content of the test questions. Instead, tests with verbal content were placed on an achievement scale and ability was measured using the sequential and simultaneous processing scales. This shift in emphasis from test content to the process needed to solve the problem put more emphasis on the cognitive activities of the examinee and resulted in a dramatic change in conceptualization of ability.

The Kaufmans' emphasis on the need for a view of ability was also important. They recognized the fact that traditional IQ lacked a theory, just as Pintner (1923) noted when he wrote "we did not start with a clear definition of general intelligence… [but] borrowed from every-day life a vague term implying all-round ability and… we [are] still attempting to define it more sharply and endow it with a stricter scientific connotation" (p. 53)." The Kaufman's stressed the point that a test of intelligence should be built upon a theory of intelligence.

Another *second-generation* ability test which was published by Naglieri and Das in 1977 further stimulated the evolution of the field. This test, called the *Cognitive Assessment System*, and the more recent CAS2 (Naglieri et al. 2014), was developed on a specific theory derived from the integration of cognitive and neuropsychological research described by A.R. Luria (1963, 1966, 1969, 1973). The theory is called PASS which stands for planning, attention, simultaneous, and successive neurocognitive abilities. Planning is the ability to perform complex decision making (related to the frontal lobes); attention is the ability to focus thinking and resist distractions (related to the brain stem); simultaneous processing ability is needed for understanding interrelationships (occipital/parietal area); and successive processing ability is used whenever sequencing is required (temporal lobes). The PASS theory was initially presented in the book *Assessment of Cognitive Processes: The PASS Theory of Intelligence* (Das et al. 1994) and elaborated by Naglieri (1999). More recently, the validity of the PASS theory as measured by the CAS and CAS2 is summarized in several resources (e.g., Das and Naglieri 2001; Naglieri 2012; Naglieri and Goldstein 2011; Naglieri and Otero 2011; Naglieri and Conway 2009; Naglieri et al. 2012, 2014).

When the KABC and CAS were introduced, these *second-generation* tests marked a change in the way intelligence was conceptualized and, just as importantly, measured. These two tests are most alike in their emphasis on measuring ability separately from academic skills. That is, they move away from the verbal/quantitative (Alpha) and nonverbal (Beta) organization of questions. The authors of these tests recognized that even though verbal and nonverbal are often described as types of intelligence, the authors of the Alpha and Beta never intended to measure two abilities, and neither did Wechsler. Wechsler's view of intelligence was not that verbal and nonverbal were two types of intelligence. Despite the fact that his tests yielded verbal and performance (nonverbal) IQ scores, Wechsler (1958) wrote: "the subtests are different measures of intelligence, not measures of different kinds of intelligence" (1958, p. 64). Boake (2002) noted that Wechsler viewed verbal and nonverbal (also labeled as performance) tests as equally valid measures of intelligence. Similarly, Naglieri (2003) clarified that the terms verbal and nonverbal "refer to the content of the test, not a type of ability" (p. 2). Moreover, Wechsler argued that nonverbal tests help to "minimize the over-diagnosing of feeble-mindness that was, he believed, caused by intelligence tests that were too verbal in content… and he viewed verbal and performance tests as equally valid measures of intelligence and criticized the labeling of performance [nonverbal] tests as measures of special abilities" (Boake 2002, p. 396).

Elimination of Army Alpha-type questions from a measure of intelligence was a bold move by these authors of *second-generation* ability tests and one that raised two important questions: (1) "How similar are verbal and quantitative test questions on an IQ test to an achievement test?" and (2) "Can verbal and quantitative tasks be taken out of a measure of intelligence without losing validity?"

Do Verbal and Quantitative Test Questions Differ from Achievement Test Questions?

It would seem reasonable that an IQ test should measure something different than an academic achievement test, but this is not the case. The verbal and quantitative portions of traditional IQ tests are remarkably similar to questions found in the achievement tests used to measure knowledge and skills. For example, verbal questions are found on both traditional IQ tests and measures of achievement. All traditional IQ tests include a measure of word knowledge just as measures of achievement do. Children are required to define a word like "bat" on subtests included in the SB-5 and WISC-IV *intelligence* tests and the WJ-III *achievement* test. The WJ-III Cognitive battery contains a Verbal Comprehension subtest that has the item like "tell me another word for small" and the WJ-III Achievement battery contains a Reading Vocabulary question: "tell me another word for little." In addition, an item on the WJ-III Reading Vocabulary achievement test is "Tell me another word for (examiner points to the word big)," and in a Cognitive battery the examiner asks something like: "Tell me another word for tiny." Additionally, the WJ-III Cognitive battery Verbal Comprehension contains Picture Vocabulary items, and the WJ-III Achievement battery includes Picture Vocabulary items some of which are the same. The WJ-III Cognitive tests also require the subject to name as many examples as possible from a given category in a 1-min time period and the same question appears on the WIAT-II Oral Expression achievement test. These examples do not comprise a complete list of item overlap but do represent the most strikingly similar questions.

Tests of intelligence and achievement also include arithmetic test questions. For example, the oldest intelligence test, now in its fifth edition is the *Stanford-Binet 5* (SB-5; Roid, 200×), contains Quantitative Reasoning items, one of which requires the child to calculate the total number of stars on a page (e.g., two stars in one box plus four in a second box plus one in a third box). Similarly, the *Wechsler Intelligence Scale for Children – Fourth Edition* (WISC-IV; Wechsler 2003) arithmetic subtest requires the child to count the number of butterflies pictured on a page. Although the scores these test items yield are used to determine the child's level of intelligence, very similar items appear on the *Wechsler Individual Achievement Test* (WIAT-II, Wechsler, 200×). On that test, for example, a Numerical Operations subtest item requires the child to determine the total number of marbles shown (e.g., 3 plus 5). Similarly, a *Woodcock-Johnson Tests of Achievement* (WJ-III; Woodcock et al. 2001) Applied Problems subtest item asks the child to count the number of crayons pictured on the stimulus book (e.g., 4). Moreover, a SB-5 Quantitative Reasoning item requires the child to complete a simple math problem (e.g., 3+2=?) just as the WJ-III Math Fluency (e.g., 5+2=?) and the WIAT-II Numerical Operations (e.g., 2+2=?) achievement tests do. There is lack of distinction between the arithmetic questions on these tests of achievement and intelligence, yet the interpretations of the scores each test yields are considerably different. In one instance the score is used to determine level of math achievement, but in the other the scores are used to determine level of intelligence. The same overlap in content is found for verbal tests.

The use of items with similar content across achievement and IQ tests is alarming for several reasons. First, because the correlation between IQ, especially verbal sections of IQ tests, and achievement test scores has been considered a source of evidence for the validity of IQ tests, the correlations between IQ tests with verbal/quantitative items and achievement tests should be considered overestimates of the relationship between ability and achievement. Moreover, the authors and/or publishers of IQ tests should justify how similar questions can be used across supposedly different constructs and how very different interpretations (e.g., achievement vs. intelligence) can be made given the similarity of item content. Second, the obvious achievement content must be justified by those that use these IQ tests when assessing culturally and/or linguistically diverse children and especially Hispanics who now constitute the largest minority group in the United States. This group is particularly at risk of being misdiagnosed because they often have parents with limited educational background and/or limited English language skills (Ramirez and de la Cruz 2002) which reduce the opportunity to acquire the knowledge of words (Hart and Risley 1995).

Can Verbal and Quantitative Tasks Be Taken Out of a Measure of Intelligence Without Losing Validity?

Explaining current academic successes and failures and predicting achievement over time is a critically important role an intelligence test can play. Having IQ test questions that measure very similar content to achievement tests enhances the predictive validity of these measures but at a cost to those with limited educational backgrounds (recall that the Beta tests were used to measure intelligence fairly for those with limited familiarity with English). The question remains, however, can *second-generation* intelligence tests correlate with achievement as well as traditional IQ? This question was examined by Naglieri (1999) who first reported that the correlations between achievement test scores with the CAS and KABC were as high as or higher than those found for the WISC-III and WJ-R. More recent findings are provided next.

Naglieri and Rojahn (2004) examined the relationships between the planning, attention, simultaneous, and successive (PASS) theory, as operationalized by the CAS, and achievement, as measured by the Woodcock-Johnson Tests of Achievement – Revised (WJ-R; Woodcock and Johnson 1989), using a nationally representative sample of 1,559 students. The correlation between the CAS Full Scale with the WJ-R Tests of Achievement was .71. More recently, Naglieri, Goldstein, DeLauder, and Schwebach (2006a) compared the Wechsler Intelligence Scale for Children – Third Edition (WISC-III; Wechsler 1991) to the CAS and the WJ-III Test of Achievement. The CAS Full Scale score correlated .83 with the WJ-III achievement scores compared to a coefficient of .63 for the WISC-III Full Scale IQ. The results suggest that when the same children took the two ability tests and those scores were correlated with the same achievement scores, both showed a strong relationship between ability and achievement, but the CAS correlated significantly higher (Naglieri et al. 2006). Most recently, Naglieri et al. (2014) reported an average correlation between the CAS2 and achievement of .70.

The KABC-II Mental Processing Index (MPI) which excludes measures of knowledge correlated, on average, .68 with total achievement on the Peabody Individual Achievement Test – Revised (PIAT-R; Markwardt 1997), .70 with total achievement on the WJ-III Achievement Scale, .74 with total achievement on the Wechsler Individual Achievement Test – Second Edition (WIAT-II; The Psychological Corporation 2001), and .74 with total achievement on the Kaufman Test of Educational Achievement – Second Edition

Comprehensive Form (KTEA-II; Kaufman and Kaufman 2004b) (Kaufman and Kaufman 2004a, Tables 8.23 to 8.30; Kaufman and Kaufman 2004b, Table 7.25). Taken as a whole, the average correlation is .69.

The studies of the CAS, CAS2, KABC, and KABC-II summarized here illustrate that a cognitive approach to understanding children's intelligence is strongly correlated with achievement test scores. Interestingly, these studies show that cognitive processes are as effective for prediction of academic performance as traditional IQ tests even though the CAS, CAS2, KABC, and KABC-II do *not* include academically laden measures such as vocabulary and arithmetic. This provides an advantage for understanding achievement strengths and weaknesses for children who come from disadvantaged environments as well as those who have had a history of academic failure.

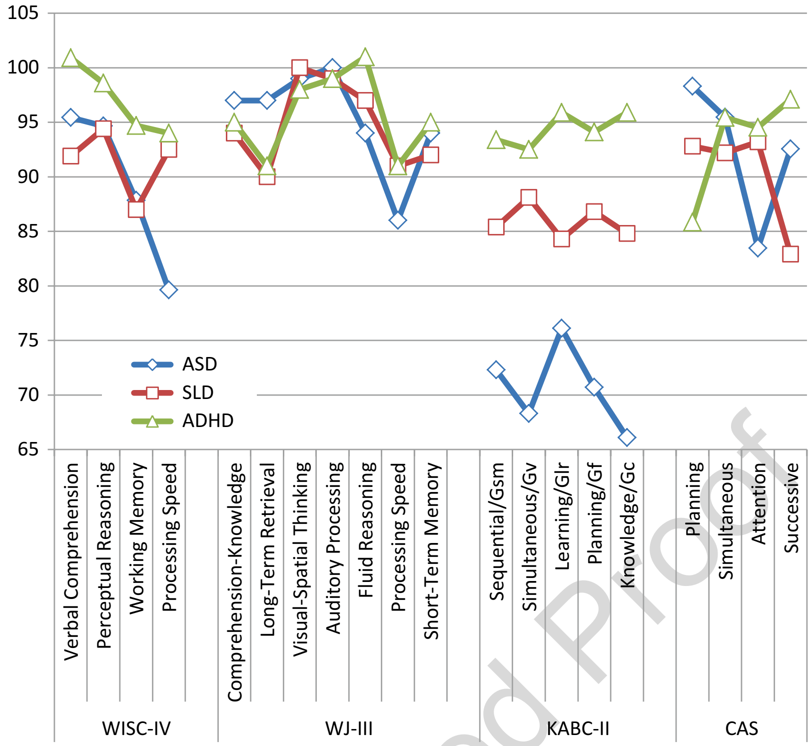
Are There Advantages to Second-Generation Intelligence Tests?

Having shown that *second-generation* ability tests correlate as well with achievement test scores as traditional IQ test which contain academic content, the next question to consider is do these tests have other advantages? For example, do these tests yield ability profiles that help understand the role a cognitive weakness may play in academic failure. A second important issue is related to fair assessment of diverse populations. More specifically, how do race differences on traditional IQ compare to those found for *second-generation* ability tests? And finally, can second-generation ability tests inform instruction and academic intervention? Each of these issues will be addressed next.

Do First- and Second-Generation Tests Detect Cognitive Problems That Underlie Academic Failure?

All intelligence tests give a full-scale score which is comprised of scales which in turn are comprised of subtests. The analysis of subtest and scale variation on tests such as the Wechsler Scales is a method called profile analysis that has been advocated by Kaufman (1994) and others (e.g., Sattler 1988) as a way to identify intellectual strengths and/or weaknesses. Information about strengths and weaknesses is then used to generate hypotheses that are integrated with other information so that decisions can be made regarding eligibility, diagnosis, and treatment. Despite the widespread use of this method, some have argued that subtest profile analysis does not provide useful information beyond that which is obtained from the IQ scores (e.g., McDermott et al. 1990; Dombrowski and Watkins 2013). Naglieri (1999) proposed that subtest analysis is problematic because of limitations in subtest reliability and validity and further suggested that what is needed is profile analysis based upon a sound theory of cognitive abilities, rather than individual subtest level analysis. Theoretically derived scales could be helpful if the ability test shows a specific pattern for a specific group of exceptional students which in turn could have implications for understanding the cognitive characteristics of the group, allow for guidance during the eligibility process (see Naglieri 2011), and guide interventions (Naglieri and Pickering 2010).

Recently, Naglieri (2011) summarized reports found in the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV; Wechsler 2003) technical manual, the Woodcock-Johnson III Tests of Cognitive Abilities (WJ-III; Woodcock et al. 2001) from Wendling et al. (2009), and CAS data from the technical manual and Naglieri, Otero, DeLauder, and Matto (2007). In the current chapter findings for students with autism spectrum disorders (ASD) from the WISC-IV (Wechsler 2003), the CAS (from Goldstein and Naglieri 2010), the WJ-III (from Wendling et al. 2009), and the KABC-II (technical manual) were added. The findings must be considered with recognition that the samples were not matched on demographic variables across the various studies, the accuracy of the diagnosis may not have been verified, and some of the sample sizes were small. Notwithstanding these limitations, the findings provide important insights into the extent to which these various tests are likely to yield scale-level profiles that are distinctive, theoretically logical, and relevant to instruction (Fig. 20.1). [AU4]



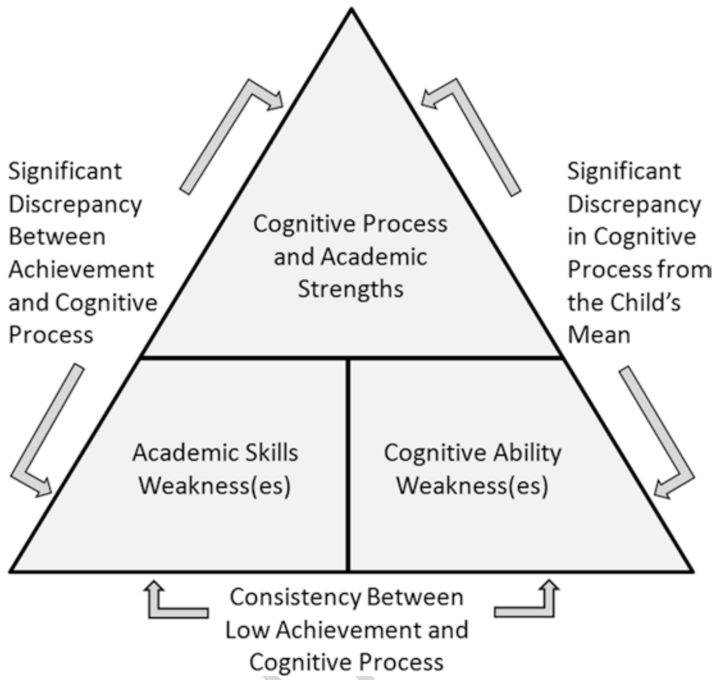
[AU5] **Fig. 20.1**

The results of the summary of scale profiles for the WISC-IV, WJ-III, KABC-II, and CAS provided in Fig. 20.2 suggest that some of these tests yield profiles that are more distinct than others across the three groups of exceptional children. The scores across all scales on the WJ-III for students with specific reading decoding difficulty (SLD) were all within the average range, and all of the KABC-II scores were in the 80s. The WISC-IV profile was lowest for the Working Memory Scale. The CAS profile showed variability across the four PASS scales with a very low score of 82.9 on the successive scale. These findings are consistent with the view that students with specific reading decoding failure also have considerable difficulty with tasks that involve sequencing of information (Das et al. 2007).

The intelligence test profiles for students with ADHD showed that all the scores for the scales on the WISC-IV, WJ-III, and KABC-II were with the average range. None of these tests provided evidence of a cognitive problem related to ADHD, except for a low score on the planning scale of the CAS. Difficulty with planning (e.g., executive function) for children with ADHD is consistent with Barkley's view that ADHD is a failure of self-control (Barkley 1997) which has been described as frontal lobe functioning (Goldberg 2009).

The results for individuals with autism spectrum disorders (ASD) show that processing speed scores on both the WISC-IV and the WJ-III were very low scores for individuals with ASD. This is similar to the findings for these two tests for individuals with ADHD. The low processing speed scores provide little insight into the cognitive characteristics of students with ASD and ADHD. Importantly, the low attention score on the CAS is consistent with the conceptualization that individuals with ASD have been described as having "difficulties in disengaging and shifting attention" (p. 214) (Klinger et al. 2009).

Fig. 20.2 Discrepancy/consistency model for specific learning disabilities



The findings for those with ASD, like the results for those with SLD and ADHD, show that the only test that had distinctive profiles for these different groups was the CAS.

Are Race Differences for Second-Generation Tests the Same as for Traditional IQ?

The need for intelligence tests that are appropriate for diverse populations of children has become progressively more important as the characteristics of the US population have changed, and recent Federal law (e.g., IDEA 2004) stipulates that assessments must be selected and administered so as to be nondiscriminatory on a racial or cultural basis. It is, therefore, critical that any measures used for evaluation be evaluated for test bias. This should include internal evidence such as reliability, item difficulty, and factor structure (see Jensen 1980) as well as mean score differences.

Some researchers have suggested that intelligence conceptualized on the basis of neuropsychological abilities is more appropriate for diverse populations (Fagan 2000; Naglieri 2005). Fagan (2000) and Suzuki and Valencia (1997) argued that measures of cognitive processes which do not rely on tests with language and quantitative content are more appropriate for assessment of culturally and linguistically diverse populations. Although there is considerable evidence for the validity of general intelligence as measured by traditional IQ tests (see Jensen 1980), researchers have traditionally found a mean difference of about 12–15 points between African-Americans and Whites on measures of IQ that include verbal, quantitative, and nonverbal tests (Kaufman and Lichtenberger 2006). Results for second-generation intelligence tests have been different.

The first evidence of smaller race differences for second-generation ability test was reported in the original KABC Manual. For children aged 2.5–12.5, without controlling for background variables, Whites (N =1,569) scored seven points higher than African-Americans (N =807) and 3 points higher than Hispanics (N=160) on the global measure of mental processing (i.e., the total test score). These differences are considerably smaller than the differences of 16 points and 11 points, respectively, reported for the WISC-R Full Scale IQ (Kaufman and Kaufman 1983, Tables 4.36 and 4.37; Kaufman et al. 2005, Table 6.7). Similar findings were reported by Naglieri (1986) in a study of 172 fifth-grade students (86 whites and 86 blacks matched on basic demographic variables) who were administered the KABC and the WISC-R. The difference between the groups on WISC-R Full Scale was 9.1 but the difference for the KABC was 6.0. Results for the KABC-II (Kaufman and Kaufman 2004) showed a similar reduction in race/ethnic differences. When controlling for gender and mother's education, African-American children at ages 3–18 years earned mean MPIs that were only five points lower than the means for White children (A. S. Kaufman and Kaufman 2004a, Tables 8.7 and 8.8; A. S. Kaufman et al. 2005, Table 6.7). Similar findings have been reported for the CAS.

Naglieri, Rojahn, Matto, and Aquilino (2005) compared PASS scores on the CAS for 298 African-American children and 1,691 White children. Controlling for key demographic variables, regression analyses showed a CAS Full Scale mean score difference of 4.8 points in favor of White children. Similarly, Naglieri, Rojahn, and Matto (2007) examined the utility of the PASS theory with Hispanic children by comparing performance on the CAS of Hispanic and White children. The study showed that the two groups differed by 6.1 points using unmatched samples, 5.1 with samples matched on basic demographic variables, and 4.8 points when demographics differences were statistically controlled. Naglieri, Otero, DeLauder, and Matto (2007) compared scores obtained on the CAS when administered in English and Spanish to bilingual children ($N=40$) referred for reading difficulties. They found a 3.0-point difference between the CAS Full Scale scores and these scores were highly correlated (.96). Otero, Gonzales, and Naglieri (2012) replicated that study with another group of students referred for reading problems and found CAS Full Scale scores that differed by less than one point and a high correlation between the scores (.94). Results for the CAS2 Full Scale scores were reported in the test manual (Naglieri et al. 2014). For children and adolescents aged 5–18 years without controlling for demographic variables, African-Americans and non–African-Americans differed by 6.3 standard scores, and with controls for demographic characteristics, the difference was 4.5. Similarly, without controlling for demographic differences, Hispanics and non-Hispanics differed on the CAS Full Scale scores by 4.5 points, and with controls for demographic characteristics, the difference was 1.8.

The importance of the findings presented above for the CAS and KABC is best understood within the context of differences found on traditional intelligence tests. Table 20.1 provides a summary of standard score differences by race for the Stanford-Binet IV (SB-IV; Roid 2003), Woodcock-Johnson Tests of Cognitive Abilities – Third Edition (WJ-III; Woodcock et al. 2001), the WISC-IV (Wechsler 2003), the KABC and KABC-II (Kaufman and Kaufman 1983, 2004), and the CAS (Naglieri and Das 1997) and CAS2

Table 20.1 Mean score differences in standard scores by race on traditional IQ and second-generation intelligence tests

Test	Difference
Traditional	
SB-IV (matched)	12.6
WISC-IV (normative sample)	11.5
WJ-III (normative sample)	10.9
WISC-IV (matched)	10.0
Second generation	
KABC (normative sample)	7.0
KABC (matched)	6.1
KABC-2 (matched)	5.0
CAS2 (normative sample)	6.3
CAS (demographic controls)	4.8
CAS2 (demographic controls)	4.3

Notes: Stanford-Binet IV (SB-IV) from Wasserman (2000); (Woodcock-Johnson III) WJ-III from Edwards and Oakland (2006); Kaufman Assessment Battery for Children (KABC) matched from Naglieri (1986); Kaufman Assessment Battery for Children – 2 from (Lichtenberger et al. 2009); CAS from Naglieri, Rojahn, Matto, and Aquilino (2005); Wechsler Intelligence Scale for Children – IV (WISC-IV) from O'Donnell (2009)

(Naglieri et al. 2014). The results for the WISC-IV are reported by O'Donnell (1009), for the SB-IV by Wasserman (2000), and the WJ-III results are from Edwards and Oakland (2006). The race differences for the KABC normative sample were reported in that test's manual (Kaufman and Kaufman 1983), and the findings for the KABC-II were summarized by Lichenberger, Sotelo-Dynega, and Kaufman (2009). Differences for the CAS were reported by Naglieri, Rojahn, Matto, and Aquilino (2005) and in the test manual for the CAS2 by Naglieri, Das, and Goldstein (2014).

The results of research on race differences for the KABC-II and CAS illustrate that second-generation tests, in contrast to traditional IQ tests, provide a more equitable way to assess diverse populations of children. The findings suggest that as a group, traditional IQ tests showed differences in ability scores between the races that are about twice as large as that found for second-generation tests. All of the traditional tests included in this table have verbal/quantitative/nonverbal content, and two of these three types of questions demand knowledge that is very similar to that required by standardized achievement tests (see Naglieri and Bornstein 2003). It is reasonable to conclude that the approach to conceptualizing and measuring intelligence taken by the authors of these second-generation ability tests resulted in smaller race difference without a loss of prediction to achievement or in the case of the CAS sensitivity to learning problems, both of which are critical components of validity.

Do Second-Generation Tests Have Relevance to Academic Intervention?

One important purpose of assessment of ability is to help decide how a student learns best and what obstacles to learning may exist. Knowing the cognitive profile of an individual student should inform instruction. This means that in addition to teaching knowledge and skills, tailored instruction should help children, for example, "to plan and control, to think and inquire, to evaluate and reflect" (Scheid 1993, p. 3). This kind of approach teaches children knowledge and skills as well as effective ways of using the abilities a student has and managing any limitations in the ability. The student is seen as an active participant who interprets information that is received, relates it to previously acquired facts, organizes and stores it for later use, develops ways of doing things, and critically examines information. Because this approach puts emphasis on both the academic skills the child must learn and the cognitive abilities the child uses in the act of learning, knowing the cognitive ability profile of a student is a critical element in a complex process that leads to effective teaching and learning (Naglieri 1999). The relationships between cognitive abilities as measured by the CAS and academic instruction have been reported in a series of research papers.

There are several resources for applying the PASS theory to academic instruction and remediation. The PASS Remedial Program (PREP; Das 1999) is an option as is the planning strategy instruction, also known as the planning facilitation method, described by Naglieri and Pickering (2003). Other resources include, for example, Kirby and Williams' (1991) Learning Problems: A Cognitive Approach, Cognitive Strategy Instruction That Really Improves Children's Academic Performance – Second Edition by Pressley and Woloshyn (1995), *Helping students become strategic learners* (Scheid 1993), and Naglieri and Pickering's (2010) book *Helping Children Learn: Intervention Handouts for use in School and Home*. The first two methods are based on empirical studies, while the remaining books contain cognitive approaches to academic interventions. The methods use structured and directed instructions based on PREP or structured but not scripted planning strategy instruction. In order to provide more details and research underlying strategy instruction and PREP, these two methods will be discussed in more detail.

Strategy Instruction. The connection between planning from PASS and interventions to improve the use of strategies has been examined in a series of studies. These investigations have involved both math and reading comprehension and have focused on the concept that children can be taught to be more strategic when they complete academic tasks and that the facilitation of plans positively impacts academic performance. The essential concept was based on the idea that *teaching* students to use strategies directly was not as advantageous as *encouraging* students to approach their work strategically. The method is designed so that children discover the value of strategy use without being specifically instructed to do so. The students are encouraged to examine the demands of the task in a strategic and organized manner. Research on this intervention method and its relationship to PASS scores on the CAS has been carefully examined in a number of important research studies.

The first two studies using planning strategy instruction showed that children's performance in math calculation improved substantially (Naglieri and Gottling 1995, 1997). The children in these two studies attended a special school for those with learning disabilities. Students completed mathematics worksheets in sessions over about a two-month period. The method designed to teach planning was applied in individual 1 on 1 tutoring sessions (Naglieri and Gottling 1995) or in the classroom by the teacher (Naglieri and Gottling 1997) two to three times per week in half hour blocks of time. Students were encouraged to recognize the need to plan and use strategies when completing mathematic problems during the intervention periods. The teachers provided probes that facilitated discussion and encouraged the students to consider various ways to be more successful. More details about the method are provided by Naglieri and Gottling (1995, 1997) and by Naglieri and Pickering (2010).

The relationship between strategy instruction and the PASS profiles for children with learning disabilities and mild mental impairments was also studied by Naglieri and Johnson (2000). The purpose of their study was to determine if children with cognitive weaknesses in each of the four PASS processes, and children with no cognitive weaknesses, showed different rates of improvement in math when given the same group planning strategy instruction. The findings from this study showed that children with a cognitive weakness in planning improved considerably over baseline rates, while those with no cognitive weakness improved only marginally. Similarly, children with cognitive weaknesses in simultaneous, successive, and attention showed substantially lower rates of improvement. The importance of this study was that the five groups of children responded very differently to the same intervention. Stated another way, the PASS processing scores were predictive of the children's response to this math intervention (Naglieri and Johnson 2000).

The effects of planning strategy instruction on reading comprehension were reported by Haddad, Garcia, Naglieri, Grimditch, McAndrews, and Eubanks (2003). This study assessed whether an instruction designed to facilitate planning would have differential benefits on reading comprehension and if improvement was related to the PASS processing scores of each child. The researchers used a sample of general education; children sorted into groups based their PASS scale profiles using the CAS. Even though the groups did not differ by CAS Full Scale scores or pretest reading comprehension scores, children with a planning weakness benefited substantially (effect size of 1.4) from the instruction designed to encourage the use of strategies and plans. In contrast, children with no PASS weakness or a successive weakness did not benefit as much (effect sizes of .52 and .06, respectively). These results further support previous research suggesting that the PASS profiles are relevant to instruction.

Iseman and Naglieri (2011) examined the effectiveness of the strategy instruction for students with LD and ADHD randomly assigned to an experimental group or a control group which received standard math instruction. They found large pre-post effect sizes for students in the experimental group (0.85), but not the control group (0.26) on classroom math worksheets, as well as standardized test score differences in Math Fluency (1.17 and .09, respectively) and Numerical Operations (.40 and −.14, respectively). One year later the experimental group continued to outperform the control group. These findings strongly suggested that students with LD and ADHD in the experimental group evidenced greater improvement in math worksheets, far transfer to standardized tests of math, and at follow-up 1 year later than the control group. The findings also illustrate the effectiveness of strategy instruction especially for those with low planning scores on the CAS.

The results of these planning strategy instruction studies using academic tasks suggest that changing the way aptitude is conceptualized (e.g., as the PASS rather than traditional IQ) and measured (using the CAS) increases the probability that an aptitude-by-treatment interaction (ATIs) is detected. Past ATI research suffered from conceptualizations of aptitudes based on the general intelligence model which did not adequately differentiate cognitive abilities which *are* related to instruction. The traditional IQ approach is very different from the PASS theory as measured by the CAS. The summary of studies provided here are particularly different from previous ATI research that found students with low general ability improve little, whereas those with high general ability respond more to instruction. In contrast, children with a weakness in one of the PASS processes (planning) benefited *more* from instruction compared to children who had no weakness or a weakness in a different PASS process. The results of these studies also suggest that the PASS profiles can help predict which children will respond to the academic instruction and which will not. This offers an important opportunity for researchers and practitioners interested in the design of instruction as suggested by Naglieri and Pickering (2003).

PREP. PREP was developed as a cognitive remedial program based on the PASS theory (Das et al. 1994). These researchers summarized [AU6] research which showed that students could be trained to use successive and simultaneous processes more efficiently, which resulted in an improvement in their performance on that process and transferred to specific reading tasks also occurred. PREP aims to improve the use of cognitive processing strategies (e.g., simultaneous and successive processes) that underlie reading. The tasks in the program teach children to focus their attention on the sequential nature of many tasks, including reading which helps the children better utilize successive processing – a very important cognitive process needed in reading decoding. PREP is also founded on the premise that the transfer of principles is best facilitated through inductive, rather than deductive, inference. The program is, therefore, structured so that tacitly acquired strategies are likely to be used in appropriate ways. For example, the tasks teach children to focus on the sequences of information included in a variety of tasks, including reading.

Support for PREP summarized elsewhere (Naglieri 2011) has shown the effectiveness of the instructional method for children with reading decoding problems. Children who received PREP in comparison to a regular reading program improved significantly on Word Attack and Word Identification tests from the Woodcock Reading Mastery Test – Revised. Learning disabled children who were randomly assigned to a PREP training and a control group that received regular classroom instruction showed significant improvement in reading decoding of real and pseudowords. When PREP was compared to a meaning-based reading program using two carefully matched groups of first-grade children, the results showed a significant improvement in reading scores for the PREP group and the gain in reading was greater than it was for the meaning-based control group. Specific relevance to the children's CAS profiles was also demonstrated by the fact that those children with a higher level of successive processing at the beginning of the program benefited the most from the PREP instruction, but those with the most improvement in the meaning-based program had higher levels of planning. Taken as a whole, these studies support the effectiveness of PREP in remediating deficient reading skills during the elementary school years and the connection between the PASS theory and intervention.

Do Second-Generation Ability Tests Aid in Determination of a Specific Learning Disability?

One of the greatest challenges to traditional IQ has been the identification of a specific learning disability, defined in IDEA 2004 as follows:

> *Specific learning disability* means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. Such term includes such conditions as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Such term does not include a learning problem that is primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

Perhaps the most essential problem with using a traditional IQ test for this purpose is the fact that alignment of the test results to the Federal regulations has been quite difficult. This is especially true because specific learning disability is defined as a disorder in basic psychology processes, but traditional IQ tests were not developed to measure basic psychological processes. Even though it is clear that the identification of a disorder in one or more of the basic psychological processes is essential for SLD eligibility determination (Fuchs and Young 2006; Hale et al. 2006; Hale et al. 2006), practitioners have been constrained by the content of traditional IQ tests which do not align with the definition.

It is logical that assessment of basic psychological processes as stipulated in IDEA 2004 be a part of any comprehensive assessment designed to determine if a child has a specific learning disability. Moreover, IDEA 2004 requires the use of a variety of assessment tools that are technically sound and nondiscriminatory to gather functional, developmental, and academic information when special education eligibility is being determined. This information should be integrated with other important findings about the child to ensure that a comprehensive evaluation is obtained. In short, documentation of a basic psychological processing disorder and academic failure is essential for SLD determination (Hale et al. 2006; Naglieri 1999, 2011), and this can be best accomplished with second-generation ability tests.

Essential to the description of a specific learning disability (SLD) is the presence of a pattern of strengths and weakness in basic psychological processes and academic skills. Of all the possible tools described in this chapter, the PASS theory as measured by the CAS2 is best suited because it yields profiles for students with SLD, works well with minority students, and has intervention implications (Haung et al. 2010; Naglieri and Otero 2011). There are three main components to eligibility determination based on this conceptualization initially presented by Naglieri (1999). First, the student has significant intraindividual *discrepancy* among the PASS scales, and the lowest PASS ability score is substantially below average. Second, there is a *discrepancy* between good PASS scores and weak achievement. Third, there is a *consistency* between poor PASS scores and academic deficits (Naglieri 1999, 2011) as illustrated in Fig. X [AU7]

An intraindividual discrepancy is examined by comparing a student's four PASS scale standard scores to determine if there exists a cognitive weakness. The purpose of these analyses is to identify PASS cognitive processing strengths (scores that are significantly greater than the student's mean score and fall above the normative average) or weaknesses (scores that are significantly lower than the student's mean score and fall below the normative average). For example, consider a student has standard scores of 114 (planning), 116 (simultaneous), 109 (attention), and 94 (successive). The successive score is 14.25 standard score points below the child's mean of 108.25 which is significant but that score of 94 is within the average range. Academic achievement scores are similar to the successive score (low portion of the average range). This would *not* be considered evidence of a disorder in one or more basic psychological processes and academic failure. In contrast, a cognitive weakness is found when, for example, a student has standard scores of 102, 104, 97, and 82 for planning, simultaneous, attention, and successive, respectively, in which case the successive score is considered a weakness and there are comparable academic achievement test scores. The successive and academic weaknesses in contrast to planning, simultaneous, attention, and academic scores that are average or higher would suggest the existence of a disorder in one or more of the basic psychological processes with academic failure as described in IDEA.

The method described above is referred to as the discrepancy/consistency model (Naglieri 1999). This method is useful for the identification of specific learning disabilities because it ensures a systematic examination of variability of both cognitive and academic achievement test scores. Determining if the cognitive processing scores differ significantly is accomplished using the method originally proposed by Davis (1959) and Silverstein (1982), popularized by Kaufman (1979), and modified by Silverstein (1993). This so-called ipsative method determines when the child's scores are reliably different from the child's average score. This technique has been applied to a number of tests including, for example, the WISC-IV (Naglieri and Paolitto 2005), the CAS (Naglieri and Das 1997), and the SB5 (Roid 2003). It is important to note that in the discrepancy/consistency model described by Naglieri (1999), the ipsative approach is applied to the PASS scales which represent four neurocognitive PASS constructs, not the subtests. This changes the method from one that relies on a clinical interpretation of the meaning of subtest variability to analysis of scales that have been theoretically defined and have higher reliability and validity. This distinction is important because the criticisms of the ipsative method (McDermott et al. 1990) have centered around subtest, not scale-level analysis.

Naglieri (1999) and Flanagan and Kaufman (2004) stressed the importance of recognizing that because a low score relative to the child's mean could still be within the average range, adding the requirement that the weakness in a processing test score is also well below average is important. In a study of PASS profiles for the CAS standardization and validity samples, Naglieri (2000) found that those students who had a PASS weakness were likely to have significantly lower achievement scores and more likely to have been identified as exceptional. That study was described by Carroll (2000) as one which illustrated what a more successful profile method could be. Davison and Kuang (2000) suggested that "adding information about the absolute level of the lowest score improves identification over what can be achieved using ipsative profile pattern information alone (p. 462)." Importantly, when Huang et al. (2010) studied PASS profiles on the CAS for large samples of students in regular education ($N=1{,}692$) and those with specific learning disabilities ($N=367$), they found ten core PASS profiles for those in regular educational and eight unique profiles from students with SLD. Huang et al. concluded that "a student with a true LD has a relatively high chance of being accurately identified when using profiles analysis on composite [PASS] scores" (p. 28). They added that their "analysis has provided evidence for the use of the PASS theory and that it appears that it has sufficient applications for diagnosis for students suspected of having a LD" (p. 28).

Section Summary

The topics covered thus far provide evidence that second-generation ability tests should be considered viable methods of evaluating children and adolescents for three important reasons. First, the KABC and CAS correlate strongly with achievement even though they do not have academic content, which suggests they have excellent validity. Second, the CAS and KABC yield small differences between Black and White (CAS and KABC and their second editions) as well as Hispanic and White (CAS, KABC-II, and CAS2) groups which provides evidence that these measures are appropriate for non-biased assessment. Third, the evidence presented shows that CAS scores reveal the weakness children with specific learning disability in reading decoding have is different from that experienced by those with other types of SLD (Haung et al. 2010) as well as ADHD and autism (Fig. 20.3).

The remainder of this chapter will focus on one of the two second-generation ability tests. Although the KABC and CAS both provide substantial advantages beyond traditional IQ, only the CAS has demonstrated specific PASS profiles for students with disabilities; it yields the smallest differences by race/ethnicity, and there is a history of research showing the relevance of PASS scores to academic instruction. For these reasons the remainder of this chapter will provide

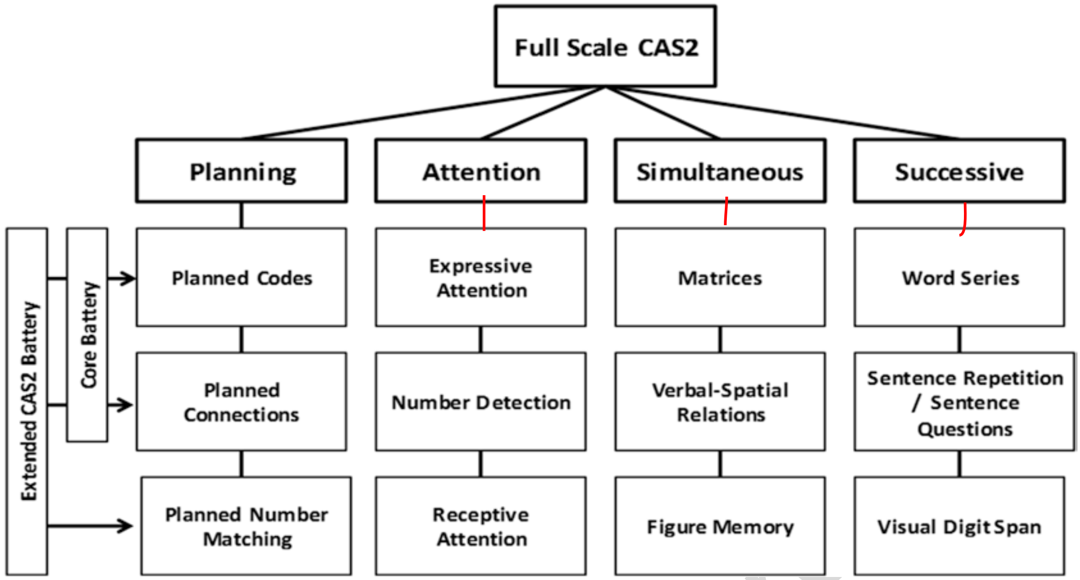


Fig. 20.3 Structure of the Cognitive Assessment System – Second Edition

a more detailed discussion of how the four PASS abilities can be measured by the CAS2.

Operationalization of the PASS Neurocognitive Abilities

The PASS theory was first operationalized by Naglieri and Das in 1997 with the publication of the CAS (and more recently the CAS2 (Naglieri et al. 2014)). Some of the research about the scores obtained from the tests developed to measure PASS was described earlier in this chapter. Although the PASS theory is more fully described elsewhere (Naglieri 1999; Naglieri and Das 1997; Naglieri and Otero 2011; Naglieri et al. 2014) and in Otero's chapter in this book, the remainder of this chapter includes a description of how each of the four PASS abilities are measured in the CAS2 and how this relates to traditional IQ.

Planning

In PASS theory, planning is a neurocognitive ability used to determine and apply strategies to solve problems and self-monitor and self-correct as needed (Naglieri et al. 2014). This includes control of actions and thoughts so that efficient solutions to problems can be achieved. Planning provides the means to solve problems for which no method or solution is immediately apparent and may involve retrieval of information as well as utilization of the other PASS abilities to process the information. Planning ability is also important when individuals reflect on events, recognizing what worked, and what did not work, and considering better problem solving in the future. The frontal lobes of the brain are directly involved in planning ability (Naglieri and Otero 2011).

The essence of tasks that measure planning is that the student must solve novel problems for which there is no previously acquired strategy and there should be minimal constraints placed on the way the student completes the task. The score a planning test yields should reflect efficiency, measured by how a student went about completing the tests and how effective the solutions were. The following tasks are used in the CAS2 to evaluate planning ability:

Planned Codes. This subtest contains four items, each with its own set of codes and particular arrangements of rows and columns. A legend at the top of each page shows which letters correspond to which codes (e.g., A, B, C, D with OX, XX, OO, XO, respectively). Just below the legend are seven rows and eight columns of letters without the codes. Children write the corresponding codes in empty boxes beneath each of the letters. The items differ in the correspondence of letters to codes and the position of the letters on the page. Students have 60 s per item to complete as many empty code boxes as possible.

Planned Connections. The Planned Connections subtest requires the student to connect numbers in sequence that appears in a quasi-random order (e.g., 1–2–3, etc.). For the more complex items, the child connects numbers and letters in sequential order, alternating between numbers and letters (e.g., 1–A–2–B, etc.). Any errors made by the child are corrected as they progresses through the task. The items are constructed so that children never complete a sequence by crossing one line over the other. This provides a means of reducing the areas to be searched when looking for the next number or letter. The subtest score is based on the total amount of time used to complete the task.

Planned Number Matching. The student's task is to find and underline two numbers that are the same in each row. Each item is composed of eight rows of numbers, with six numbers per row. Two of the six numbers in each row are the same. The length of numbers differs on the various rows. Numbers increase in digit length from one digit on the first row of item 1 to seven digits on the eighth row of item 4. There are four rows for each digit length and a total of four pages of numbers. Children aged 5–7 are provided an example followed by two test items. Ages 8–18 are provided an example and two test items. Each row of numbers was carefully developed to maximize benefits of strategy usage in the identification of correct matches. This approach resulted in items with some rows that contain numbers that start with unique numbers, some rows that include numbers with similar digit strings, and some rows that contain numbers that end with similar numbers.

Attention

Attention is a neurocognitive ability used to selectively focus on a particular stimulus while inhibiting responses to competing stimuli presented over time (Naglieri et al. 2014). Attention is a basic component of intelligent behavior involving allocation of resources and effort. Arousal, attention, effort, and capacity are concepts that have a complex relationship and importance for understanding behavior. Luria stated that optimal conditions of arousal are needed before the more complex forms of attention involving "selective recognition of a particular stimulus and inhibition of responses to irrelevant stimuli" (Luria 1973, p. 271) can occur. Attention is conceptualized as a mental activity that provides focused, selective cognitive activity over time and resistance to distraction. The process is involved when a person must demonstrate focused, selective, sustained, and effortful activity. The longer the attention needed, the more the activity necessitates vigilance. Intentions and goals mandated by the planning process control attention, whereas knowledge and skills play an integral part in the process as well.

Tasks that measure attention include target and nontarget stimuli that are multidimensional with the requirement that the person has to identify one aspect of the target (e.g., the color blue) and resist responding to distractions (e.g., a word red written in blue ink) as in the Stoop test (Lezak 1995). This kind of a task requires selective focus of attention over time, an ability that is necessary for learning to take place. The following tasks are used in the CAS2 to evaluate attention:

Expressive Attention. The Expressive Attention subtest consists of two age-related sets of three items. Students ages 5–7 years are presented with three items consisting of seven rows that each contain six pictures of common animals, with each picture depicted as either big (1 in. by 1 in.) or small (1/2 in. by 1/2 in.). In each of three items, the student is required to identify whether the animal depicted is big or small in real life, ignoring the relative size of the picture on the page. In item 1, the pictures are all the same size.

In item 2, the pictures are sized appropriately (i.e., big animals are depicted with big pictures, and small animals are depicted with small pictures). In item 3, the realistic size of the animal often differs from its printed size. Students ages 8–18 years are presented with three items consisting of eight rows of five words each. In item 1, students are asked to read four black-and-white color words (blue, yellow, green, and red) that are presented in random order. In item 2, students are asked to name the colors of four-colored rectangles (printed in blue, yellow, green, and red) that are presented in random order. In item 3, the four-color words are printed in a different-color ink than the color-word name and are presented in random order. In this item, students are required to name the color of the ink in which the word is printed rather than read the word.

Number Detection. Each Number Detection item presents the student with a page of approximately 200 numbers. Students are required to underline specific numbers (ages 5–7 years) or specific numbers in a particular font (ages 8–18 years) on a page with many distractors. There are four pages of numbers, each of which is scored for the number correct, number of false detections, and time.

Receptive Attention. The Receptive Attention subtest consists of four-item sets, each containing 60 picture pairs (ages 5–7 years) or 180 letter pairs (8–18 years). Both versions require the student to underline pairs of objects or letters that either are identical in appearance or are the same from a lexical perspective (i.e., they have the same name). There are four pages of numbers, each of which is scored for the number correct, number of false detections, and time.

Simultaneous

Simultaneous is a neurocognitive ability used to integrate separate stimuli into a single whole or interrelated group (Naglieri et al. 2014). The essence of simultaneous processing is that separate elements must be combined into a conceptual whole. This ability is involved in visual-spatial tasks as well as those language activities that require comprehensive of grammatical structures. The spatial aspect of simultaneous ability involves both the perception of stimuli as a group or whole and the formation of visual images. The grammatical dimension of simultaneous processing allows for the integration of words into ideas through the comprehension of word relationships, prepositions, and inflections, so the person can obtain meaning.

Tasks designed to measure simultaneous processing often have visual-spatial content. One well-known measure of simultaneous processing is progressive matrices. Traditional intelligence tests often include subtests that use the progressive matrix format, as do many nonverbal intelligence tests such as the Naglieri Nonverbal Ability Test (Naglieri 2011). These tests are often categorized as perceptual reasoning or nonverbal, but from PASS, matrices measure simultaneous ability. This ability can also be measured using verbal content which requires comprehension of the grammatical components of language such as comprehension of word relationships and understanding of prepositions and inflections (Naglieri 1999). The Verbal-Spatial Relations subtest on the CAS is an example of this type of a subtest (Naglieri et al. 2014). This arrangement of subtests allows for measurement of simultaneous ability across verbal and nonverbal contents. The tests used to evaluate simultaneous neurocognitive ability on the CAS2 are as follows:

Matrices. Matrices are a multiple-choice subtest that utilizes shapes and geometric elements that are interrelated through spatial or logical organization. Students are required to analyze the relationship among the parts of the item and solve for the missing part by choosing the best of five options. The raw score is the total number of items correctly answered.

Verbal-Spatial Relations. Verbal-Spatial Relations is a multiple-choice subtest in which each item consists of six drawings and a printed question at the bottom of each page. The examiner reads the question aloud, and the child is required to select the option that matches the verbal description. The items require the evaluation of logical grammatical relationships (e.g., "which picture shows a ball in a basket under a table?"), which demands simultaneous processing with verbal content. The raw score is the total number of items correctly answered.

Figure Memory. For each Figure Memory item, the examiner presents the student with a two- or three-dimensional geometric figure for 5 s. The picture is then removed, and the student is presented with a response page that contains the original figure embedded in a large, more complex geometric pattern. The student is required to trace the original figure with a red pencil in the Figure Memory Response Form. The raw score is the total number of items correctly answered.

Successive

Successive is a neurocognitive ability used to work with information that is arranged in a specific serial order where each part follows the other in a strictly defined order (Naglieri et al. 2014). Successive processing is involved in the perception of stimuli in sequence as well as the formation of sounds and movements into a specific order. This type of ability is necessary for the recall of information in order as well as phonological analysis and the syntax of language (Das et al. 1994). Deficits with successive processing are also associated with early reading problems in young children, as it requires a child to learn sounds in a sequential order.

Tasks used to measure successive processing include digit span forward (as well as the recall of numbers, words, or hand movements) which is found on many tests of ability. These tests are sometimes described as measures of working memory or sequential processing (a concept very close to successive processing in PASS theory). Sometimes a backwards version is included which involves successive as well as planning processing abilities (Schofield and Ashman 1987). The successive tasks included in the CAS and CAS2 provide a way to measure this ability using tests that demand repeating a sentence using the correct series of words (Sentence Repetition) as well as comprehension of sentences that are understood only by appreciating the sequence of words (Sentence Questions). Additionally, CAS2 has a visual digit span test, allowing for measurement of successive processing across auditory and visual modalities. The tests used in the CAS2 to measure successive processing ability are as follows:

Word Series. The Word Series subtest utilizes nine single-syllable, high-frequency words: book, car, cow, dog, girl, key, man, shoe, and wall. The examiner reads aloud a series of two to nine of these words at the rate of one word per second. The student is required to repeat the words in the same order as stated by the examiner. The raw score is the total number of items correctly answered.

Sentence Repetition. The Sentence Repetition subtest (administered only to ages 5–7 years) requires the student to repeat syntactically correct sentences containing little meaning, such as "The blue is yellowing." The raw score is the total number of items correctly answered.

Sentence Questions. The Sentence Questions subtest (administered only to ages 8–18 years) requires the student to listen to sentences that are syntactically correct but contain little meaning and answer questions about the sentences. For example, the student is read the sentence "The blue is yellowing" and then asked the following question: "Who is yellowing?" The raw score is the total number of items correctly answered.

Visual Digit Span. Visual Digit Span subtest requires the student to recall a series of numbers in the order in which they were shown using the Stimulus Book. Each item that is 2–5 digits in length is exposed for the same number of seconds as there are digits. Items with six digits or more are all exposed for a maximum of 5 s. The raw score is the total number of items correctly answered.

The CAS2 subtests described above can be combined into an 8-subtest Core Battery or a

12-subtest Extended Battery to yield four scores following the PASS theory: planning, attention, simultaneous, and successive, and a total score called the Full Scale. The subtests are all individually administered tests designed explicitly to yield scores to evaluate the four PASS neurocognitive abilities for children and adolescents aged 5 years 0 months through 18 years 11 months. It was normed on a representative sample of 1,342 students. The test manual provides a complete summary of reliability and validity of the CAS2 as well as interpretive and intervention information. See Naglieri et al. (2014) for more details.

Closing Thoughts

The purpose of this chapter was to organize 100 years of progress in the area of IQ tests, twentieth-century traditional ideas about intelligence, and a second-generation of intelligence tests. The essence of the discussion has been about the tools and concepts we have used in this most important field of applied psychology. I have argued that there are several important issues that need to be recognized and will be reiterated here.

First, traditional IQ began July 20, 1917, with the development of the verbal (quantitative) and nonverbal IQ test format. This format has dominated the IQ testing industry since that time and has been used in all individual- and group-administered IQ tests.

Second, the tests developed for the US Army were designed to test many recruits in the shortest amount of time and with the least amount of effort needed for scoring. There was no theory of intelligence that guided the selection or development the Army Alpha and Beta tests. These tests have been accepted as measures of intelligence and in fact the IQ score has become synonymous with the term intelligence.

Third, traditional IQ tests include questions that are very similar to tests found in achievement tests, especially, for example, vocabulary, word analogies, and math word problems. The role of knowledge needed to answer these types of questions was recognized as undesirable by the original authors of the Army Alpha. Despite the fact

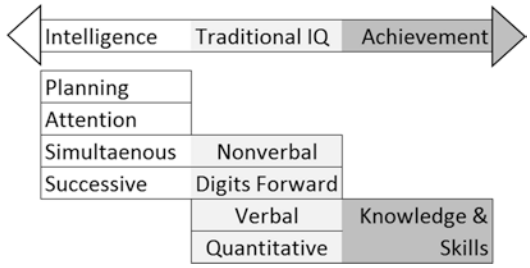


Fig. 20.4 Intelligence achievement continuum

that such tests are often indistinguishable from measures of *achievement*, users of traditional IQ tests have ignored this problem and compounded the issue by calling such tests measures of *verbal intelligence*.

Fourth, about 80 years after the birth of traditional IQ, a second-generation of intelligence tests appeared. The first was the KABC (Kaufman and Kaufman 1983) and second was the CAS (Naglieri and Das 1997). These tests were designed with a conceptualization of intelligence (KABC) or a specific theory of intelligence (CAS); and importantly, they did not include the verbal and arithmetic test items found in traditional IQ. On a continuum from pure intelligence to pure achievement, the second-generation tests were clearly distinct from tests of achievement and traditional IQ (see Fig. 20.4). Traditional IQ tests share some overlap with second-generation IQ tests (mainly the nonverbal portion of traditional IQ tests and simultaneous scales of the CAS and KABC). The verbal and quantitative portions of traditional IQ share overlap with achievement tests. The academic content of traditional IQ tests poses considerable problem for test validity and assessment of diverse populations as well as those with learning difficulties.

Fifth, research has clearly supported second-generation intelligence tests over traditional IQ. Newer tests offer several advantages including (a) a theory that can be used to create scales on a test that represent a psychological construct, (b) greater fairness to minorities and to those with limited academic skills, (c) scores that represent different abilities according to the theory upon which the test was developed, (d) greater ability to identify special populations of individuals with intellectual disorders related to behavioral or academic disabilities, and (e) enhanced ability to link second-generation intelligence test scores to interventions.

The challenge faced by second-generation intelligence tests, despite their clear advantages over traditional IQ, is inertia. Traditional IQ has 100 years of use and acceptance and countless numbers of research studies and books written about them and their interpretation. But as Neil deGrasse Tyson, author of the new guide to the Cosmos, recently commented (2014) on the value of traditional wisdom "In practically ever idea we have as humans, the older version of it is *not* better than the new version" (p. 80). It is time for the field of intelligence testing to embrace new ideas of what intelligence may be and how best to measure it.

References [AU8] [AU9]

Barkley, R. A. (1997). *ADHD and the nature of self-control*. New York: Guilford.

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.

Carroll, J. B. (2000). Commentary on profile analysis. *School Psychology Quarterly, 15*, 449–456.

Das, J. P. (1999). *PASS reading enhancement program*. Deal: Sarka Educational Resources.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes*. Boston: Allyn & Bacon.

Das, J. P., Janzen, T., & Georgiou, G. K. (2007). Correlates of Canadian native children's reading performance: From cognitive styles to cognitive processes. *Journal of School Psychology, 45*, 589–602.

Davis, F. B. (1959). Interpretation of differences among averages and individual test scores. *Journal of Educational Psychology, 50*, 162–170.

Davison, M. L., & Kuang, H. (2000). Profile patterns: Research and professional interpretation. *School Psychology Quarterly, 15*, 457–464.

Edwards, O. W., & Oakland, T. D. (2006). Factorial invariance of Woodcock-Johnson III scores for African Americans and Caucasians Americans. *Journal of Psychoeducational Assessment, 24*, 358–366.

Fagan, J. R. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law, 6*, 168–179.

Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. Hoboken: Wiley.

Fuchs, D., & Young, C. L. (2006). On the irrelevance of intelligence in predicting responsiveness to reading instruction. *Exceptional Children, 73*, 8–30.

Goldberg, E. (2009). *The new executive brain: Frontal lobes in a complex world*. New York: Oxford University Press.

Haddad, F. A., Garcia, Y. E., Naglieri, J. A., Grimditch, M., McAndrews, A., & Eubanks, J. (2003). Planning facilitation and reading comprehension: Instructional relevance of the PASS theory. *Journal of Psychoeducational Assessment, 21*, 282–289.

Hale, J. B., Kaufman, A. S., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA: Using RTI and cognitive assessment methods. *Psychology in the Schools, 43*, 753–770. [AU10]

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American Children*. Baltimore: Brookes.

Haung, L. V., Bardos, A. N., & D'Amato, R. C. (2010). Identifying students with learning disabilities: Composite profile analysis using the cognitive assessment system. *Journal of Psychoeducational Assessment, 28*, 19–30.

IDEA. (2004). Individuals with Disabilities Education Improvement Act of 2004 (P.L. 108–446).

Iseman, J., & Naglieri, J. A. (2011). A cognitive strategy instruction to improve math calculation for children with ADHD: A randomized controlled study. *Journal of Learning Disabilities, 44*, 184–195.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. Circle Pines: American Guidance.

Kaufman, A. S., & Kaufman, N. L. (Eds.). (2001). *Learning disabilities: Psychological assessment and evaluation*. Cambridge, UK: Cambridge University Press.

Kaufman, A. S., & Kaufman, N. L. (2004). Kaufman assessment battery for children second edition. Circle Pines, MN: American Guidance.

Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., & Kaufman, N. L. (2005). *Essentials of KABC-II assessment*. New York: Wiley.

Klinger, L., O'Kelly, S. E., & Mussey, J. L. (2009). Assessment of intellectual functioning in Autism spectrum disorders. In S. Goldstein, J. A. Naglieri, & S. Ozonoff (Eds.), *The assessment of autism spectrum disorders* (pp. 209–252). New York: Guilford.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York: Oxford University Press.

Lichenberger, Sotelo-Dynega, & Kaufman (2009). [AU11]

Lichtenberger, E. O., Sotelo-Dynega, M., & Kaufman, A. S. (2009). The Kaufman assessment battery for children. In J. A. Naglieri & S. Goldstein (Eds.), *A practitioner's guide to assessment of intelligence and achievement* (2nd ed., pp. 61–93). New York: Wiley. [AU12]

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York: Basic Books.

Matarazzo, J. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore: Williams & Wilkins.

McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment, 8*(3), 290–302.

Naglieri, J. A. (1999). *Essentials of CAS assessment*. New York: Wiley.

Naglieri, J. A. (2000). Can profile analysis of ability test scores work? An illustration using the PASS theory and CAS with an unselected cohort. *School Psychology Quarterly, 15*, 419–433.

Naglieri, J. A. (2011). The discrepancy/consistency approach to SLD identification using the PASS theory. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 145–172). Hoboken: Wiley.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment, 21*, 244–260.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive assessment system*. Itasca: Riverside Publishing Company.

Naglieri, J. A., & Goldstein, S. (2011). Assessment of cognitive and neuropsychological processes. In S. Goldstein, J. A. Naglieri, & M. DeVries (Eds.), *Understanding and managing learning disabilities and ADHD in adolescence and adulthood* (2nd ed., pp. 137–160). New York: Wiley.

Naglieri, J. A., & Gottling, S. H. (1995). A cognitive education approach to math instruction for the learning disabled: An individual study. *Psychological Reports, 76*, 1343–1354.

Naglieri, J. A., & Gottling, S. H. (1997). Mathematics instruction and PASS cognitive processes: An intervention study. *Journal of Learning Disabilities, 30*, 513–520.

Naglieri, J. A., & Johnson, D. (2000). Effectiveness of a cognitive strategy intervention to improve math calculation based on the PASS theory. *Journal of Learning Disabilities, 33*, 591–597.

Naglieri, J. A., & Otero, T. (2011). Cognitive assessment system: Redefining intelligence from a neuropsychological perspective. In A. Davis (Ed.), *Handbook of pediatric neuropsychology* (pp. 320–333). New York: Springer.

Naglieri, J. A., & Paolitto, A. W. (2005). Ipsative comparisons of WISC-IV index scores. *Applied Neuropsychology, 12*, 208–211.

Naglieri, J. A., & Pickering, E. (2010). *Helping children learn: Intervention handouts for use in school and at home* (2nd ed.). Baltimore: Brookes.

Naglieri, J. A., & Rojahn, J. R. (2004). Validity of the PASS theory and CAS: Correlations with achievement. *Journal of Educational Psychology, 96*,174–181.

Naglieri, J. A., Rojahn, J. R., Matto, H. C., & Aquilino, S. A. (2005). Black white differences in intelligence: A study of the PASS theory and cognitive assessment system. *Journal of Psychoeducational Assessment, 23*,146–160.

Naglieri, J. A., Goldstein, S., Delauder, B. Y., & Schwebach, A. (2006a). WISC-III and CAS: Which correlates higher with achievement for a clinical sample? *School Psychology Quarterly, 21*, 62–76.

Naglieri, J. A., Rojahn, J., Matto, H. (2006b). Hispanic and non-Hispanic children's performance on PASS cognitive processes and achievement. *Intelligence* [AU13]

Naglieri, J. A., Otero, T., DeLauder, B., & Matto, H. (2007). Bilingual Hispanic children's performance on the English and Spanish versions of the cognitive assessment system. *School Psychology Quarterly, 22*, 432–448.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014). *Cognitive assessment system* (2nd ed.). Austin: Pro-Ed.

O'Donnell, L. (2009). The Wechsler intelligence scale for children. In J. A. Naglieri & S. Goldstein (Eds.), *A practitioner's guide to assessment of intelligence and achievement* (4th ed., pp. 153–190). New York: Wiley.

Otis, A. S. (1918). *An absolute point scale for the group measurement of intelligence*. https://ia600401.us.archive.org/28/items/cu31924073896650/cu31924073896650.pdf. Reprinted from the Journal of Educational Psychology, IX.

Pintner, R. (1923). *Intelligence testing*. New York: Henry Holt.

Pressley, M. P., & Woloshyn, V. (1995). *Cognitive strategy instruction that really improves children's academic performance* (2nd ed.). Cambridge, MA: Brookline.

Ramirez, R. R., & de la Cruz, G. P. (2002). *The Hispanic population in the United States: March 2002* (Current population reports, P20-545). Washington, DC: U.S. Census Bureau.

Roid, G. (2003). *Stanford-Binet* (5th ed.). Itasca: Riverside.

Scheid, K. (1993). *Helping students become strategic learners*. Cambridge, MA: Brookline.

Schofield, N. J., & Ashman, A. F. (1987). The cognitive processing of gifted, high average, and low average ability students. *British Journal of Educational Psychology, 57*, 9–20.

Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education, 1*, 1–57.

Silverstein, A. B. (1982). Pattern analysis as simultaneous statistical inference. *Journal of Consulting and Clinical Psychology, 50*, 234–240.

Silverstein, A. B. (1993). Type I, Type II, and other types of errors in pattern analysis. *Psychological Assessment, 5*, 72–74.

Suzuki, L. A., & Valencia, R. R. (1997). Race-ethnicity and measured intelligence. *American Psychologist, 52*, 1103–1114.

Wechsler, D. (1992). *Wechsler individual achievement scale*. San Antonio: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio: The Psychological Corporation.

Wendling, B. J., Mather, N., & Shrank, F. A. (2009). Woodcock-Johnson III tests of cognitive abilities. In J. A. Naglieri & S. Goldstein (Eds.), *A practitioner's guide to assessment of intelligence and achievement* (pp. 191–232). New York: Wiley.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of cognitive abilities*. Itasca: Riverside Publishing Company.

Author Queries

Chapter No.: 20 0002192606

Queries	Details Required	Author's Response
AU1	Please provide details of Yerkes (1920, 1921), Wechsler-Bellevue Intelligence Scale (1939), Yoakum and Yerkes (1920), Otis-Lennon (1979), Wechsler (2014), Kaufman (1994), Prifitera and Saklofske (1998), Weiss et al. (2006), Naglieri and Kaufman (2008), Luria (1963, 1966, 1969), Das and Naglieri (2001), Naglieri (1986, 2012), Naglieri and Conway (2009), Wechsler (1958, 1991), Naglieri (2003), Woodcock and Johnson (1989), Markwardt (1997), The Psychological Corporation (2001), Kaufman and Kaufman (2004a, b), Sattler (1988), Dombrowski and Watkins (2013), Goldstein and Naglieri (2010), Naglieri (2005), Kaufman and Lichtenberger (2006), Naglieri et al. (2007), Otero et al. (2012), Naglieri and Pickering (2003), Kirby and Williams' (1991), Wasserman (2000) in the reference list.	
AU2	Please provide organization name for Jack A.Naglieri.	
AU3	Please provide opening quotes in the sentence "They recognized the fact that traditional…".	
AU4	Please check if inserted citations for Figs. 20.1 and 20.3 are okay.	
AU5	Please provide caption for Fig. 20.1.	
AU6	Please check latter part of sentence starting "These researchers summarized…" for sense.	
AU7	Please provide appropriate citation for Fig. X.	
AU8	Please cite Kaufman and Kaufman (2001), Lezak et al. (2012), Matarazzo (1972), Naglieri et al. (2006b), Otis (1918), Share and Stanovich (1995) and Wechsler (1992) in text.	
AU9	References Wendling et al. (2009), Naglieri (2011), Naglieri and Otero (2011), Kaufman and Kaufman (1983, 2004) are one and the same. Hence duplicate entries have been deleted. Please check.	
AU10	Please confirm the inserted volume number for Hale et al. (2006).	
AU11	Please update the reference for Lichenberger et al. (2009.)	
AU12	Please check if initial inserted for Lichenberger et al. (2009) is okay.	
AU13	Please provide volume number and page range for Naglieri et al. (2006b).	