School Psychology Review,
2001, Volume 30, No.1

What Does the Cognitive Assessment System (CAS) Measure? Joint Confirmatory Factor Analysis of the CAS and the Woodcock-Johnson Tests of Cognitive Ability (3rd Edition)

Timothy Z. Keith
Alfred University

John H. Kranzler
University of Florida

Dawn P. Flanagan
St. John's University

Abstract. Results of recent research by Kranzler and Keith (1999) raised important questions concerning the construct validity of the Cognitive Assessment System (CAS; Naglieri & Das, 1997), a new test of intelligence based on the planning, attention, simultaneous, and sequential (PASS) processes theory of human cognition. Their results indicated that the CAS lacks structural fidelity, leading them to hypothesize that the CAS Scales are better understood from the perspective of Cattell-Horn-Carroll (CHC) theory as measures of *psychometric g, processing speed, short-term memory span,* and *fluid intelligence/broad visualization.* To further examine the constructs measured by the CAS, this study reports the results of the first joint confirmatory factor analysis (CFA) of the CAS and a test of intelligence designed to measure the broad cognitive abilities of CHC theory—the Woodcock-Johnson Tests of Cognitive Abilities-3rd Edition (WJ III; Woodcock, McGrew, & Mather, 2001). In this study, 155 general education students between 8 and 11 years of age (M = 9.81) were administered the CAS and the WJ III. A series of joint CFA models was examined from both the PASS and the CHC theoretical perspectives to determine the nature of the constructs measured by the CAS. Results of these analyses do not support the construct validity of the CAS as a measure of the PASS processes. These results, therefore, question the utility of the CAS in practical settings for differential diagnosis and intervention planning. Moreover, results of this study and other independent investigations of the factor structure of preliminary batteries of PASS tasks and the CAS challenge the viability of the PASS model as a theory of individual differences in intelligence.

Naglieri and Das (1997) developed the Cognitive Assessment System (CAS) to assess the planning, attention, and simultaneous-successive (PASS) cognitive processes of children and adolescents from 5 through 17 years of age. According to the PASS theory of intelligence,

Data collection for this research was supported by a grant from Measurement/Learning/Consultants LLC; we are grateful for their support. We solely are responsible for any errors and for the opinions expressed.

Correspondence should be sent to Timothy Z. Keith, Division of School Psychology, Alfred University, Saxon Drive, Alfred, NY 14802. Email: tzkeith@alfred.edu

information processing is related to three functional units of the brain: (a) *planning* entails the formulation, selection, and regulation of plans of action; (b) *attention* involves the distribution of cognitive resources and effort; and (c) *simultaneous* and *successive* processing comprise the cognitive processes used in the acquisition, storage, and retrieval of information (for a summary, see Das, Naglieri, & Kirby, 1994). Naglieri (1997) stated that "the PASS processes are dynamic in nature, respond to the cultural experiences of the individual, are subject to developmental changes, and form an interrelated (correlated) interdependent system" (p. 250). Although the PASS processes are seen to be related, they are nonetheless conceptualized as "physiologically and functionally distinct" (Naglieri, Das, & Jarman, 1990, p. 429).

At the current time, the CAS is one of the few tests of intelligence derived from a theory of information processing and the only test based entirely on the PASS theory (for reviews, see Flanagan, Genshaft, & Harrison, 1997). "The purpose of the CAS is to measure specific abilities defined as PASS cognitive processes. These processes are considered the basic dimensions of ability" (Naglieri, 1999a, p. 10). Naglieri and Das claim that the CAS has substantiated validity for the following uses: "diagnosis of learning strengths and weaknesses; classification (learning disabilities, attention deficit, mental retardation, giftedness); eligibility decisions (meeting state or federal criteria); and consideration of the appropriateness of particular treatment, instructional, or remedial programs" (p. 9; cf. Naglieri, 1999a).

Results of recent research, however, raised serious questions about the construct validity of the CAS (Keith & Kranzler, 1999; Kranzler & Keith, 1999; Kranzler, Keith, & Flanagan, 2000). Kranzler and Keith (1999) analyzed the standardization data of the CAS with confirmatory factor analysis (CFA) techniques to address several important and unresolved issues suggested by research on preliminary batteries of PASS tasks (Carroll, 1995; Kranzler & Weng, 1995a, 1995b; Naglieri, Das, Stevens, & Ledbetter, 1991). Although results of their multisample CFA indicated that the CAS measures the same constructs across its 12-year age span, the model reflecting the implied hierarchical structure of the CAS provided a poor fit to the data. They also found that the average correlation between factors reflecting planning and attention across age groups exceeded +.90, indicating that the planning and attention processes on the CAS are difficult to distinguish. In addition, three of the four factors underlying the PASS Scales were found to have inadequate specificity to support their interpretation. Given that only the Successive Scale had enough uniqueness to be interpreted alone, ipsative analysis of the PASS Scales on the CAS is ill-advised. Finally, the correlated PASS model—the theoretical model held by Naglieri and Das to underlie the CAS—did not provide the best fit to the data. Consistent with independent research on preliminary batteries of experimental CAS (i.e., PASS) tasks by Carroll (1995) and by Kranzler and Weng (1995a), the model that provided the best fit to the data was a third-order hierarchical model, with one general factor (i.e., psychometric *g*) at the apex of the factor hierarchy, one combined Planning/Attention factor at an intermediate level, and four first-order factors corresponding to the PASS processes. The relations among the factors for this best-fitting model cannot be accounted for by PASS theory, but are quite consistent with Carroll's (1993) three-stratum theory (for further discussion, see Keith & Kranzler, 1999; Kranzler & Keith, 1999; Naglieri, 1999b). Recently, Kranzler et al. (2000) replicated these substantive conclusions in an independent examination of the factor structure of the CAS.

Taken as a whole, results of Kranzler and colleagues' analyses revealed that the CAS lacks *structural fidelity*, indicating that the relations among the scaled scores of the CAS (i.e., subtests, PASS Scales, and Full Scale) are not consistent with the theory upon which the test is based (viz., PASS theory). Without structural fidelity, the construct validity of the CAS simply cannot be established. Kranzler and colleagues' CFA results, therefore, do not support the use of the CAS in practical settings for differential diagnosis or for planning educational interventions based on the PASS Scales.

What Does the CAS Measure?

If the CAS does not measure the constructs that Naglieri and Das intended it to measure, then what does it measure? Theories of the structure of cognitive abilities, which attempt to explain the organization, or structure, of individual differences in cognitive abilities, are extremely useful for understanding the abilities measured by new tests of intelligence such as the CAS. At the present time, the Cattell-Horn-Carroll (CHC) theory of cognitive abilities, which incorporates Cattell-Horn's *Gf-Gc* theory (e.g., Horn, 1994) and Carroll's (1993, 1997) three-stratum theory, is widely regarded as the best description of the structure of human cognitive abilities. In the CHC theory, cognitive abilities are classified at three levels of generality. At Stra-

Table 1
Summary of Findings About the CAS

We Know:

- The CAS measures the same attributes for children between 5-7 years.
- The CAS measures the same attributes for children and youth between 8-17 years.
- Of the *non-hierarchical* models examined, the PASS model provides the best fit.
- The constructs measured by the CAS are overlapping and related.
- The Planning and Attention factors are virtually indistinguishable.
- A three-level hierarchical model explains the CAS structure better than does the correlated PASS model.
- The first-order PASS factors are, in part, explained by a general factor of cognitive ability, psychometric *g*.
- The Planning and Attention factors are largely explained by another, less-general factor in addition to *g*.
- The Planning, Simultaneous, and Attention factors have very little unique variance; scale scores representing these constructs should not be interpreted in isolation.
- Scores representing Successive and Planning/Attention constructs may be interpreted in isolation.

We Believe:

- The Planning and Attention tests are primarily measures of processing speed or perceptual speed.
- The Simultaneous Scale is a mixture of fluid intelligence (*Gf*) and broad visualization (*Gv*) rather than simultaneous mental processing.
- The Successive Scale is a measure of short-term memory span rather than of successive mental processing.
- The Planning tests measure little in common beyond processing speed.
- The Attention tests measure little in common beyond processing speed.

From "Independent Confirmatory Factor Analysis of the Cognitive Assessment System (CAS): What Does the CAS Measure?" by J. H. Kranzler & T. Z. Keith, 1999, *School Psychology Review, 28*, p. 139. Copyright 1999 by the National Association of School Psychologists. Reprinted with permission.

tum III, the apex of the factor hierarchy, is psychometric *g*, the general factor shared by all tests of cognitive ability. Stratum II reflects the 10 broad cognitive abilities specified in *Gf-Gc* theory (e.g., fluid intelligence, crystallized intelligence, processing speed). Stratum I consists of approximately 70 relatively narrow abilities (e.g., inductive reasoning, memory span, rate-of-test taking). CHC theory is extremely useful to test developers, because it specifies the number and kinds of abilities to assess (Daniel, 1997).

In Table 6 of their article, Kranzler and Keith (1999) summarized the results of their study on the structure of the CAS (shown here in Table 1). This table is divided into sections labeled "We Know" and "We Believe." The first section consists of conclusions that were supported by the empirical results of their study. Because the evidence supporting these conclusions was strong, Kranzler and Keith were confident that other researchers would arrive at the same conclusions through examination of their findings or through their own independent investigations of the constructs underlying the CAS (e.g., Kranzler et al.,2000). The second section of Table 1, labeled "We Believe," consists of hypotheses about the underlying structure of the CAS. Based on their CFA results and on inspection of CAS task demands, Kranzler and Keith asserted that the CAS tests are best viewed from the perspective of Carroll's (1993, 1997) three-stratum theory of human cognitive abilities (now known as CHC theory) as measures of *processing speed* (rather than planning and attention), *short-term memory span* (rather than successive coding), and a mixture of *fluid intelligence* and *broad visualization* (rather than simultaneous coding) (cf. McGrew, 1997; McGrew & Flanagan, 1998). They further speculated that the substantial psychometric *g* that underlies the CAS is not necessarily a good *g*. They stated that, "because the CAS predominantly measures a rather narrow range of abilities that are not central aspects of psychometric *g* (viz., memory span and perceptual speed), the best estimate of *g* on the CAS appears to be the Simultaneous Scale" (p. 139).

Despite having empirical results and contemporary theory to support their hypotheses, Kranzler and Keith (1999) cautioned that, "although we maintain that most independent researchers would support these conclusions, we recognize that other explanations are possible. These latter conclusions, therefore, are best viewed as hypotheses about the structure of the CAS that require substantiation by future studies" (p. 140). They added, however, that many of their hypotheses could be tested via conjoint CFA of the CAS tests with another, better-understood test of intelligence. As Carroll (1995) noted, the factor analyses conducted by Naglieri, Das, and their colleagues "always have been very limited, with no more than perhaps 10 or 12 tests in any one study, in such a way that it is difficult to define or cross-identify the factors found" (p. 400). He further stated that they:

> Have concentrated on a small number of tests that they claim define concepts derived from Luria's theories, but they have not demonstrated that these tests consistently measure these concepts, and *only* [emphasis in the original] those concepts. They have made few attempts to examine relationships between their tests and tests that have been used in more than 50 years of research in cognitive abilities, nor have they considered adequately the possibility that their PASS tests measure dimensions of ability, such as *g*, fluid intelligence, crystallized intelligence, spatial ability, perceptual speed, and many others, that have been recognized in cognitive ability research. (p. 408)

Although Naglieri, Das, and their colleagues have published the results of CFAs of the small battery of CAS tests purported to measure the PASS processes (Naglieri & Das, 1997)[1], they have not examined the possible relations between the CAS factors and other widely recognized factors of human cognitive ability (for reviews, see Carroll, 1993, 1997).

Aims of the Present Study

It is unclear whether the CAS tests and scales measure the constructs intended by the test authors—Planning, Attention, and Simultaneous and Successive mental processing—or whether the tests and scales of the CAS measure several constructs from CHC theory, as argued by Kranzler and Keith (1999) and others (Carroll, 1993, 1997; Flanagan, McGrew, & Ortiz, 2000; Keith & Kranzler, 1999; McGrew & Flanagan, 1998). The purpose of this study was to test these competing theoretical explanations of the constructs measured by the CAS. To do so, we conducted a series of joint CFAs of the CAS tests with a test designed to measure the constructs from CHC theory (the most recent version of the Woodcock-Johnson Psycho-Educational Battery, the WJ III; Woodcock et al., 2001). As Keith and Kranzler (1999) noted, a joint CFA of the CAS and the WJ III will shed much needed light onto the constructs measured by the CAS.

Briefly, we used these competing theoretical orientations to develop a series of predictions concerning the results of joint CFAs of the CAS and the WJ III. For each analysis, we used CHC theory to make a prediction concerning the results of that analysis, and we also used PASS theory to make a prediction concerning the results of the analysis. The results from each analysis supported one theory or the other. Our CHC predictions about the outcome of these analyses were derived from Kranzler and Keith (1999) and Keith and Kranzler (1999) and are consistent with the CHC classifications of the CAS tests offered by Flanagan et al. (2000) and by McGrew and Flanagan (1998). Our predictions about the outcome of the analyses from the PASS perspective are based on the writings of Naglieri, Das, and colleagues (e.g., Das et al., 1994; Naglieri, 1999a, 1999b; Naglieri & Das, 1997).[2] These predictions are explicated in the Results section as we discuss each analysis.

Method

Participants

Participants were 155 students in Grades 3 to 6 (59 boys, 96 girls), from the general education classes of elementary schools in North Central Florida and New York City. The age of participants ranged from 8 to 11 years (M = 9.81 years, SD = .88). None of the participants was receiving special education or related services. The racial group breakdown of the sample was as follows: 73 African American, 66 Caucasian, and 12 Asian American children, with four participants of unreported race. Nineteen of the 155 participants were of Hispanic descent (regardless of race). The primary language of 86% of participants was English, with another 11% coming from bilingual homes. All participants were treated in accordance with the "Ethical Principles of Psychologists and Code of Conduct" (American Psychological Association, 1992).

Instruments

Cognitive Assessment System (CAS). The CAS was developed to assess the PASS cognitive processes of children and adolescents (Naglieri & Das, 1997). The standard CAS battery consists of 12 subtests. The PASS processes are reflected in four scales and their respective subtests:

- *Planning* (Matching Numbers, Planned Codes, and Planned Connections);
- *Attention* (Expressive Attention, Receptive Attention, and Number Detection);
- *Simultaneous* (Nonverbal Matrices, Figure Memory, and Verbal-Spatial Relations); and
- *Successive* (Word Series, Sentence Repetition, and either Speech Rate or Sentence Questions, depending on the age of the individual; children between 5 to 7 years of age are administered Speech Rate, whereas 8- to 17-year-olds are given Sentence Questions).

PASS Scale scores are based on an equally weighted composite of the subtests underlying each respective scale. The Full Scale (FS) score is based on an equally weighted aggregate of the PASS subtests and is interpreted as an estimate of overall cognitive functioning. Further information on the PASS theory, organization of the scales, and development of subtests can be found in the *Interpretative Handbook* (Naglieri & Das, 1997, pp. 1-25). Additional information on the PASS theory and the CAS can be found in Das et al. (1994) and in Naglieri (1999a).

Woodcock-Johnson Tests of Cognitive Abilities-3rd Edition (WJ III). The Woodcock-Johnson Tests of Cognitive Abilities—Revised (WJ-R; Woodcock & Johnson, 1989) was developed to assess the broad cognitive abilities represented in contemporary CHC theory, and is the only intelligence test currently in use to do so. The WJ-R has strong research support as measures of the CHC constructs (e.g., Bickley, Keith & Wolfle, 1995; Keith, 1997; McGrew & Flanagan, 1998), and preliminary evidence suggests that the newest version of the WJ-R, the WJ III, provides even more complete measures of CHC abilities (McGrew & Woodcock, 2001). The WJ III, therefore, is the best available choice for these joint CFAs designed to determine whether the CAS measures the PASS or the CHC constructs.

The WJ III was developed to assess the broad cognitive abilities of CHC theory of individuals between 2 and 90+ years of age. The standard WJ III cognitive battery consists of seven standard and three supplemental tests. The extended WJ III battery includes 10 additional measures. The broad abilities measured in the present study were assessed with the following WJ III standard (S) and extended (E) tests, as well as a number of WJ-R/special research (R) tests that were used during the standardization of the WJ III:

- *Short-Term/Working Memory (Gsm)*: Memory for Sentences (R), Memory for Words (E), and Auditory Working Memory (S);
- *Processing Speed (Gs):* Visual Matching (S), Cross Out (R), and Decision Speed (E);
- *Comprehension-Knowledge (Gc):* Verbal Comprehension (Picture Vocabulary + Oral Vocabulary + Analogies; for the present research, however, a composite of Picture Vocabulary and Oral Vocabulary was used) (S), General Information (E), and Story Recall (a *Gc* test on the Achievement battery);
- *Visual-Spatial Thinking (Gv):* Block Rotation (R) and Visual Closure (R);
- *Auditory Processing (Ga):* Sound Blending (S) and Incomplete Words (S);
- *Long Term Retrieval (Glr):* Visual-Auditory Learning (S) and Memory for Names (R); and
- *Fluid Reasoning (Gf):* Concept Formation (S), Analysis-Synthesis (E), and Numerical Reasoning (R).
- In addition, the Planning (*Gv;* E) and the Auditory Attention (*Ga;* E) tests were used in several analyses.

The General Intellectual Ability (GIA) score of the WJ III is based on a differentially weighted composite of the cognitive tests, consisting of either the standard tests only or the standard and extended tests combined. The GIA is designed to provide an estimate of general intelligence (*g*) or overall cognitive functioning. The Brief Intellectual Ability cluster, a composite of three equally weighted tests, is also available. The CHC Cluster scores are based on an equally weighted composite of the tests underlying each respective CHC ability. These scores are used to identify cognitive ability strengths and weaknesses (Woodcock et al., 2001). Finally, seven standard tests can be combined into differentially weighted scholastic aptitude composites for predicting achievement in different achievement domains. Further information on the CHC theory, organization of WJ III clusters, and the development and validity of these tests can be found in the WJ III Technical Manual (McGrew & Woodcock, 2001).

It is important to note that this research was not designed to assess the validity of the WJ III. Instead, we have used the WJ III to provide markers for the CHC abilities in order to determine whether the CAS measures the PASS or the CHC constructs. For this reason, the WJ tests used in this research were chosen to represent the CHC constructs rather than to constitute a typical WJ battery.

Procedure

Consent forms were sent to parents/guardians of all children in Grades 3 to 5 of the general education classes of several elementary schools in North Central Florida and New York City. Participants in this study included children who returned signed consent forms. The CAS and WJ III were administered individually by trained examiners under standardized conditions in a counterbalanced design. Examiners were advanced graduate students in school psychology programs, all of whom had successfully completed a graduate-level seminar and practicum in intellectual assessment. The tests were administered in one or two sessions. Total testing time was approximately 2.5 to 3.5 hours.

Statistical Analyses

We conducted several sets of CFAs to test the hypotheses in this research. Those analyses are described briefly here, and in more detail as results are presented. In the first set of analyses, we conducted a series of *conjoint* first-order CFAs that included the intended factor structure of the CAS and WJ III, respectively. For this set of analyses, we examined the correlations among the factors across the two tests to determine whether they conformed to predictions based on CHC theory or predictions based on PASS theory. Second, we conducted a series of conjoint *hierarchical* CFAs to investigate whether the respective first-order factors derived from the CAS and the WJ III are a reflection of higher-order factors in the CHC theory (e.g., *Gs*, *Gsm*). Third, in a series of *integrated* CFAs, we tested the viability of including the CAS tests on specific WJ III factors based on CHC theory. Fourth, we tested several models derived from PASS theory. These included models in which WJ III tests were loaded on specific factors based on PASS theory, as well as models that included several additional tests that should behave differently depending on whether the CAS is measuring the PASS processes or the broad cognitive abilities in CHC theory. Fifth, and finally, we examined the relation between the *g* factor underlying the CAS and the WJ III.

The primary focus of this research was to compare competing CFA models rather than evaluate the fit of a single model, in isolation. For that reason, we emphasized fit indices that are useful for comparing models rather than "stand-alone" fit indices (for a discussion of the difference, see Boomsma, 2000, or Kranzler & Keith, 1999). In particular, the change in chi-squared ($\Delta\chi^2$), along with degrees of freedom and associated probability, was used for a statistical comparison of competing, nested models. Although χ^2 has been generally abandoned as a stand-alone measure of fit, $\Delta\chi^2$ is quite useful for comparing nested models, especially when sample size is not excessive (Keith, 1997). The Akaike Information Criterion (AIC) was also used to compare models (Boomsma, 2000). Because the AIC has the advantage of not requiring that the models be nested, it was used to compare nonnested, competing models. Stand-alone fit indices are reported only for the initial model.

The Amos 4.0 computer program was used to conduct the CFAs reported in this research (Arbuckle & Wothke, 1999). Maximum-likelihood estimation of age-corrected raw scores was used for all analyses. To create age-corrected raw scores, raw (CAS) or W (WJ) scores were regressed on age in months, with the residuals representing age-corrected raw scores. Such scores should provide more complete age correction than standard scores (Jensen & Sinha, 1993). The specific procedures for each step in the analyses are described in more detail along with the results of the analyses (for further information on this method of CFA, see Keith, 1997).

Results

Conjoint CFAs

Figure 1 shows the basic model used for the first series of analyses. The right side of the model shows the subtests from the CAS, and the factors Naglieri and Das (1997) intended them to measure. The factor names derived from both the PASS and CHC theories are included in this figure. The left side of the figure shows the basic structure of the WJ III, including both the tests used in these analyses and the factors they are intended to measure. Of primary interest are the curved lines between the CAS and the WJ III factors, which represent the correlations among latent constructs. For each analysis, we first predict the results from the perspective of the CHC and PASS theories. We then report the results of the unconstrained correlations between factors to determine the adequacy of the predictions based on each theory. Finally, we further evaluate these predictions by constraining the fac-

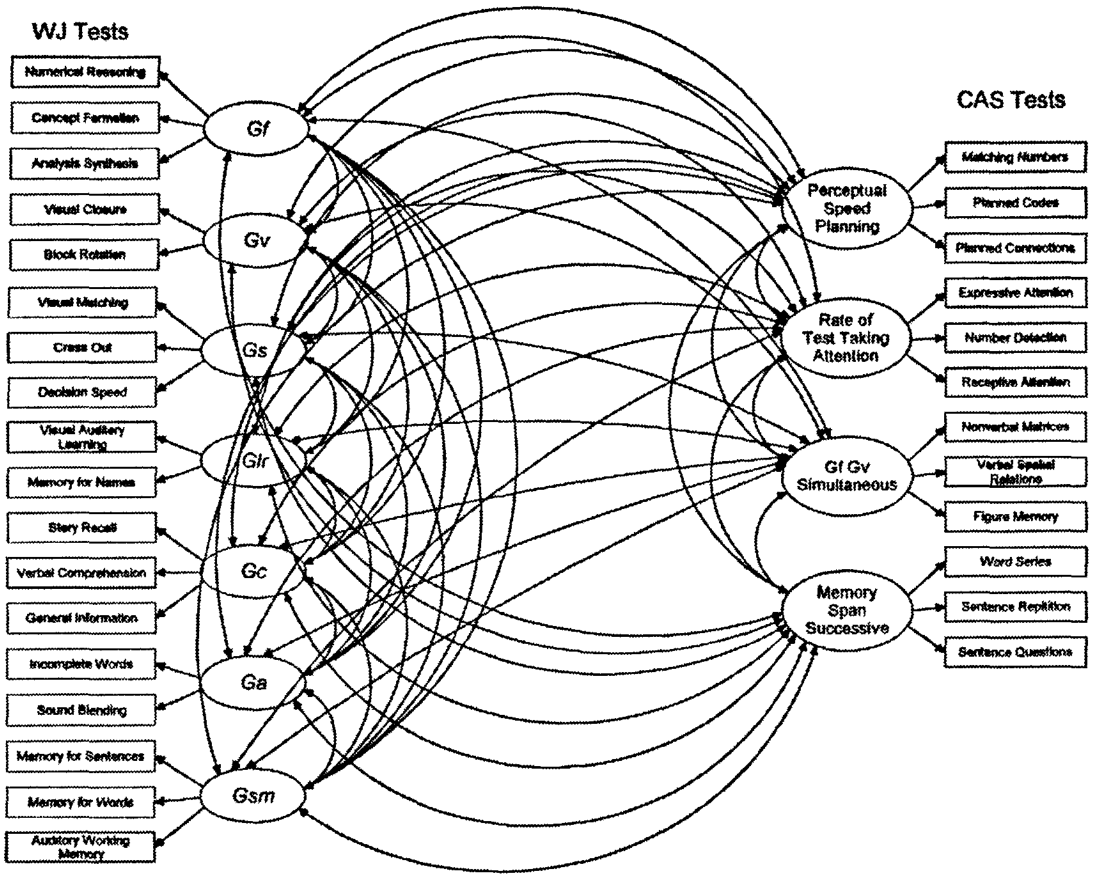

Figure 1. The unconstrained, first-order, conjoint model. The CAS subtests and factors are shown on the right, and the WJ III tests and factors are shown on the left. The curved lines connecting the factors represent correlations among latent factors. The unique and error variances of the measured variables are not shown to improve the readability of the figure.

tor correlations to certain values based on CHC or PASS theory and examining the change in model fit following the imposition of those constraints.

Do Planning and Attention tests measure processing speed? Kranzler and Keith (1999) hypothesized that the Planning and Attention factors on the CAS measure processing speed. Likewise, Flanagan and colleagues classified the Planning and Attention tests of the CAS as measures of *Gs* (Flanagan et al., 2000; McGrew & Flanagan, 1998). From the perspective of CHC theory, one would predict that the CAS Planning and Attention factors from Figure 1 will correlate highly with the *Gs* factor on the WJ III. In contrast, from the perspective of the PASS theory, high correlations between these factors should not exist. As Naglieri stated in a recent interview with Joseph (1999), "the suggestion that Planning and Attention scales are measures of processing speed is simply not supported by theory nor by research" (p. 8; see also Das et al., 1994; Naglieri, 1999a, 1999b; Naglieri & Das, 1997). Based on PASS theory, it seems reasonable to predict that correlations between these factors will be no higher than those with any other factor. When the factors were allowed to correlate freely (as in Figure 1), the correlation between *Gs* and Planning was .98 and between *Gs* and Attention was .88. Results of these unconstrained models, therefore, support the prediction derived from CHC theory, and do not support the prediction derived from PASS theory.

We then constrained the correlation between *Gs* and the Planning and Attention factors to .490, which was the average correlation among the WJ III factors. These constraints operationalize the prediction, from the perspective of the PASS theory, that the *Gs* factor will correlate no more highly with the Planning and Attention factors than with any other factor. The imposition of these constraints, however, led to a statistically significant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 43.259$, $df = 2$, $p < .001$), suggesting that these additional constraints should be rejected.[3] Information pertaining to the fit of this model is shown in Table 2.

Finally, we constrained the correlation between *Gs* and the Planning and Attention factors to 1.00 and compared the fit of this model to that of the unconstrained model. Imposition of these constraints represents an extremely strict and demanding interpretation of Kranzler and Keith's (1999) hypothesis, because it specifies that the Planning and Attention factors are identical to the *Gs* factor. Nevertheless, this model provided an equivalent

Table 2
Comparison of the Conjoint CFA Models Embodying Competing Hypotheses About the Nature of the CAS

Model	$\Delta\chi^2$ (*df*)	*p*	AIC
Unconstrained conjoint model	475.317 (350)		705.317
Correlation of Planning, Attention with *Gs* = .49	43.259(2)	<.001	744.576
Correlation of Planning, Attention with *Gs* = 1.00	4.902(2)	.086	706.219
Correlation of Successive with *Gsm* = 65	25.520(1)	<.001	728.837
Correlation of Successive with *Gsm* = 1.00	.797(1)	.372	704.114
Planning & Attention = *Gs*, and Successive = *Gsm*	5.701(3)	.127	705 018

Note. Each model was compared to the unconstrained conjoint model. The $\Delta\chi^2$ value reported for the unconstrained model is the χ^2. Other stand-alone fit indices for this unconstrained model were: Root Mean Square Error of Approximation = .048 (90% CI = .037 - .059), TLI = .907, CFI = .925.

fit to the data as the unconstrained model ($\Delta\chi^2 = 4.902$, $df = 2$, $p = .086$). Because the constrained model is the more parsimonious of the two, it is the superior model. In sum, results of this series of analyses support the predictions derived from CHC theory, and fail to support those derived from PASS theory. CHC theory, in turn, suggests that the Planning and Attention tests from the CAS measure processing speed rather than planning and attention skills.

The finding of the indistinguishability of the *Gs* with the Planning and Attention factors may be interpreted beyond its support for CHC theory, as well. What does this indistinguishability mean? At the most basic level, the finding that the correlations between the Gs and the Planning and Attention factors are indistinguishable from 1.0 means that whatever the underlying source of variance (i.e., mechanism, skill, or ability) that causes examinees to differ in their performance on the *Gs* factor is the same as that which causes them to differ in their performance on the Planning and Attention factors. All valid variation is co-variation. The factors may require other skills, but those skills are not a source of variation in examinees' performance on those factors. In other words, the Planning and Attention factors measure the same underlying construct as does the *Gs* factor. The fact that the CHC-derived prediction was supported, whereas the PASS prediction was rejected, argues that the construct measured by these three factors is processing speed, *Gs*, not planning or attention.

Do Successive Processing tests measure short-term memory? Kranzler and Keith (1999) also hypothesized that the Successive factor on the CAS measures short-term memory and Flanagan and colleagues classified the Successive tests of the CAS as measures of *Gsm* (e.g., McGrew & Flanagan, 1998). From the perspective of CHC theory, one would predict that the Successive factor on the CAS should correlate highly with the *Gsm* factor on the WJ III. Although Das et al. (1994) contended that most tests of cognitive ability contain poor measures of successive processing (p. 126), they did conjecture that "the successive component may relate to a *limited* [emphasis added] extent with WJ-R Short-Term Memory (*Gsm*)" (p. 127). According to PASS theory, therefore, one presumably would not predict a high correlation between these factors. Rather, based on their comments, it seems reasonable to predict that the CAS Successive factor will correlate somewhat more highly with *Gsm* than the average correlation among factors, but perhaps not extremely highly. Results of the unconstrained model, however, revealed that the correlation between the CAS Successive factor and the WJ III *Gsm* factor was not significantly different from 1.00,[4] which means that these two factors are statistically indistinguishable.

Because PASS theory suggests that the Successive factor may be related to a limited extent to *Gsm*, we constrained the correlation between these two factors to .65, a value somewhat greater than the mean interfactor correlation. This constraint is consistent with the hypothesis that the correlation between the Successive and *Gsm* factors should be higher than that between the Successive factor and other factors, but not at a level approaching unity. This constraint, however, resulted in a statistically significant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 25.520$, $df = 1$, $p < .001$), suggesting that it should be rejected.

In contrast, when the correlation between the Successive and *Gsm* factors was constrained to 1.00—consistent with the CHC theory-derived hypothesis that the Successive factor measures short-term memory—the fit was equivalent to the unconstrained model. Thus, for this series of analyses, the CHC-derived predictions were supported, whereas the PASS-derived predictions were not, a finding that supports the CAS Successive tests as measures of short-term memory rather than successive mental processing.

The final row of Table 2 shows the fit of the model in which all the accepted constraints were made. Again, this model, in which the correlations of Planning and Attention with *Gs* and Successive with *Gsm* were constrained to 1.00, was equivalent to, but more parsimonious than the original, unconstrained model. In sum, results of this series of conjoint CFAs support the predictions from CHC but fail to support those from PASS theory. These analyses also support the hypotheses that the Planning and Attention factors of the CAS measure processing speed and that the Successive factor of the CAS measures short-term memory.

Do Simultaneous Processing tests measure both fluid intelligence and broad visualization? Kranzler and Keith (1999) hypothesized that the Simultaneous factor on the CAS measures a mixture of *Gf* and *Gv*. This hypothesis was consistent with the CHC classifications of the CAS Simultaneous tasks provided by McGrew and Flanagan (1998). From the perspective of CHC theory, therefore, one would predict that the Simultaneous factor will correlate at a fairly high level with both the *Gf* and *Gv* factors, but perhaps not too highly. According to PASS theory, in comparison, one would predict a relatively high correlation with *Gv only*. As Das et al. (1994) stated, "the WJ-R Visual (*Gv*) factor will likely be similar to our [S]imultaneous factor" (p. 127). To the best of our knowledge, Naglieri and Das have not discussed the possible relation between *Gf* and the Simultaneous factor. Thus, CHC theory and PASS theory do not yield distinctly different predictions for the Simultaneous tests. Results of the unconstrained model revealed a correlation of .68 between the Simultaneous and *Gv* factors and .77 between the Simultaneous and *Gf* factors. These results are consistent with Kranzler and Keith's (1999) hypothesis about the constructs measured by the CAS Simultaneous Scale, but they also are consistent with the partial prediction based on the PASS theory. Results of these analyses, therefore, could be taken as support for either the PASS or CHC theoretical perspectives. Because these two positions did not lead to distinctly different predictions, no constraints were imposed on this conjoint model.

Hierarchical Conjoint CFAs

For the second set of analyses, several first-order factors from the conjoint CFAs were specified as indicators of higher-order CHC factors. In the first hierarchical conjoint CFA, the WJ III *Gs* factor and the CAS Planning and Attention factors were specified as indicators of a higher-order *Gs* factor. From the perspective of CHC theory, the Planning and Attention factors should have substantial loadings on this second-order *Gs* factor, and forcing them on the same factor as the *Gs* factor from the WJ III should not result in a significant degradation in the fit of this "Hierarchical *Gs*"

Table 3
Comparison of the Hierarchical Conjoint CFA Models Testing CHC Theory-Derived Hypotheses About the Nature of the CAS

Model	$\Delta\chi^2$ (*df*)	*p*	AIC
Unconstrained, nonhierarchical conjoint model	475.317 (350)[a]		705.317
Hierarchical *Gs* factor	18.991 (16)[b]	.269	692.308
Hierarchical *Gs* and *Gsm* factors	7.648 (6)[c]	.265	687.956
Hierarchical G*s*, *Gsm*, *Gf* and *Gv* factors	1.980 (3)[d]	.577	683.936
Hierarchical *Gs*, *Gsm*, and *Gf* factors	15.581 (5)[d]	.008	693.537
Hierarchical *Gs*, *Gsm*, and *Gv* factors	2.110 (5)[d]	.834	680.066

[a]The $\Delta\chi^2$ shown for the unconstrained, nonhierarchical conjoint model is the χ^2. It is the same baseline model as in Table 2. [b]Model compared to the unconstrained nonhierarchical model. [c]Model compared to the Hierarchical *Gs* factor model. [d]Model compared to the Hierarchical *Gs* and *Gsm* factors model.

model. In contrast, according to PASS theory, because the Planning and Attention factors do not measure *Gs*, forcing these three first-order factors on the second-order *Gs* factor should lead to a statistically significant degradation in the fit of this model. Table 3 shows the fit of the Hierarchical *Gs* model in comparison to the unconstrained conjoint model (i.e., the model used as the baseline model in the previous comparisons). As shown in the table, imposition of these additional constraints did not result in a statistically significant degradation in fit (i.e., increase in the $\Delta\chi^2$) over the baseline conjoint model ($\Delta\chi^2 = 18.991$, $df = 16$, $p = .269$). Results of this hierarchical model support the predictions made from CHC theory and fail to support those made from PASS theory. These results, therefore, provide further support for the hypothesis that the CAS Planning and Attention factors measure processing speed.

In the second hierarchical conjoint CFA, the *Gsm* factor on the WJ III and the Successive factor on the CAS were both loaded on a hierarchical *Gsm* factor. Based on CHC theory, one would predict that this "Hierarchical *Gsm*" model will provide an equivalent fit to the data as did the previous model (i.e., the Hierarchical *Gs* model). According to PASS theory, in comparison, specification of the Successive and *Gsm* factors as reflections of the same underlying cognitive ability will result in a significant degradation in model fit. As shown in Table 3, this model modification resulted in a small, statistically nonsignificant increase in $\Delta\chi^2$ over the previous model ($\Delta\chi^2 = 7.648$, $df = 6$, $p = .265$). These findings, therefore, support the hypothesis that the CAS Successive factor measures short-term memory.

We conducted the three final analyses in this set of hierarchical conjoint CFAs to further investigate the Simultaneous factor on the CAS. We examined the relative fit of models in which the Simultaneous factor was loaded on: (a) both a hierarchical *Gf* factor and a hierarchical *Gv* factor, (b) only a hierarchical *Gf* factor, and (c) only a hierarchical *Gv* factor. Because Kranzler and Keith (1999) hypothesized that the Simultaneous factor on the CAS is a mixture of *Gf/Gv*, the superior fit of the first model would provide the most support for their perspective. Although Naglieri, Das, and their colleagues have not addressed the possible relation between the Simultaneous factor of the CAS and *Gf*, they did suggest that the Simultaneous factor will be "similar" to *Gv*. The superior fit of the third model, therefore, presumably provides the most support for their perspective (we are assuming that "similar to" can be interpreted as "measures the same construct as"). Still, the predictions from the two perspectives do not differ greatly.

As shown in Table 3, in the first of these models, the addition of hierarchical *Gf* and *Gv* factors resulted in no significant degradation of model fit ($\Delta\chi^2 = 1.980$, $df = 3$, $p = .577$). One parameter in the model was clearly impossible, however, suggesting that this model should be rejected. In the second model, placing the CAS Simultaneous factor on a hierarchical *Gf* factor resulted in acceptable parameter estimates, but a significant degradation in model fit over the *Gs* and *Gsm* model ($\Delta\chi^2 = 15.581$, $df = 5$, $p = .008$). The third and final model, in which the CAS Simultaneous factor was placed on a hierarchical *Gv* factor, provided the best fit of the three. The $\Delta\chi^2$ over the *Gs* and *Gsm* model was not statistically significant ($\Delta\chi^2 = 2.110$, $df = 5$, $p = .834$), the AIC was lower than that for the *Gf* and *Gv* or the *Gf*-only model, and the parameter estimates were reasonable.[5] These analyses, therefore, offer partial support for both the CHC- and PASS-derived predictions; those predictions, however, are not sufficiently different to allow a convincing test. These analyses do not support the hypothesis that the CAS Simultaneous factor is a mixture of *Gf* and *Gv*; rather, they support the hypothesis that the CAS Simultaneous factor is related to *Gv* only. It is important to note that acceptance of the latter hypothesis is not inconsistent with the CHC theory per se and does not necessarily support the PASS theory over the CHC theory. Further testing will be needed to determine whether this hierarchical factor should be considered *Gv*, a mixture of *Gv/Gf* (from a CHC perspective), or Simultaneous (from a PASS perspective). The final model, incorporating all accepted changes, is shown in Figure 2.

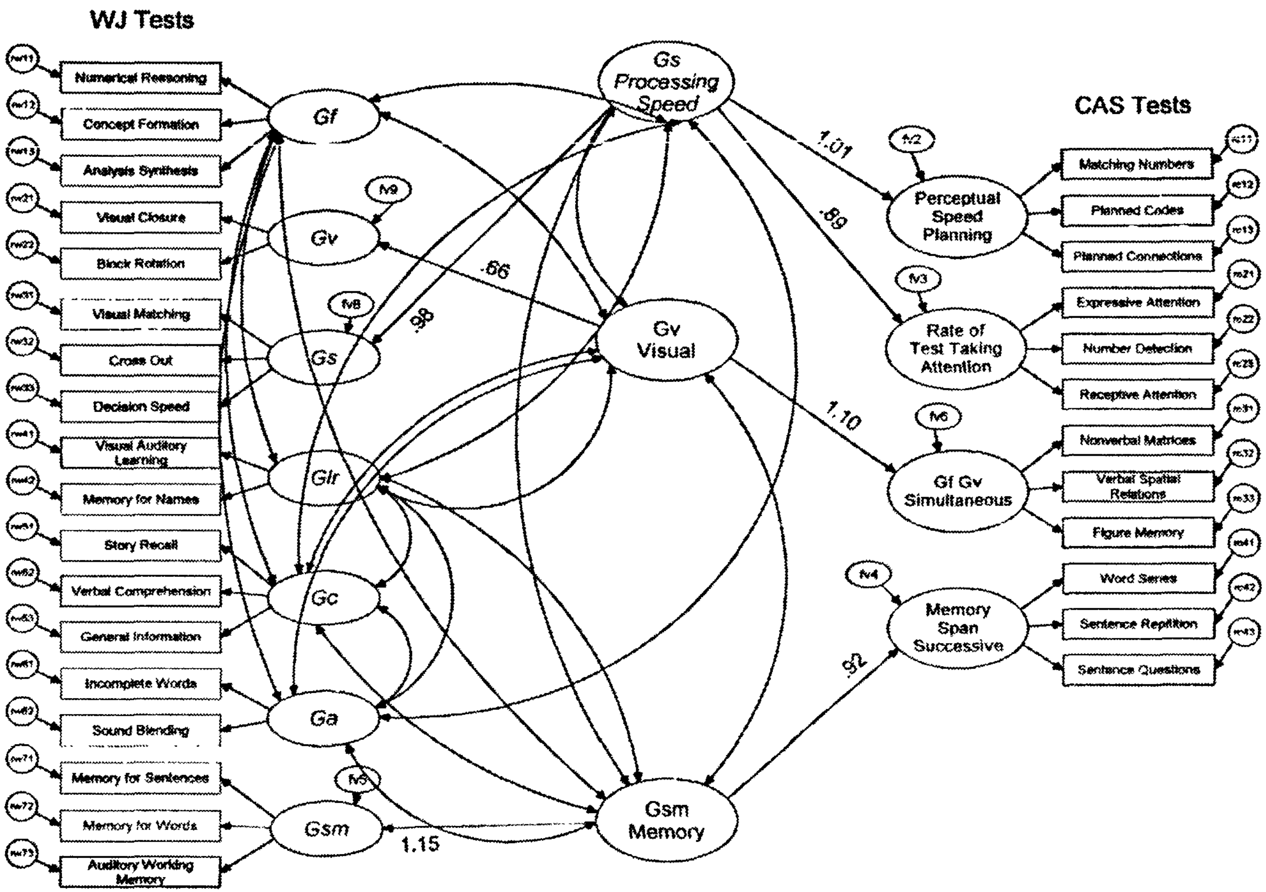

Figure 2. The final hierarchical, conjoint model. The values for the factor intercorrelations and factor loadings are not shown to improve readability. The small circles labeled rw and rc represent the unique and error variances (residuals) of the WJ III and the CAS tests, respectively.

Integrated CHC Models

For the third set of CFAs, we examined the relative fit of models after loading the CAS subtests on certain a priori specified WJ III factors based on CHC theory. For example, for the first such *integrated* CFA, the Planning and Attention tests were loaded on the WJ III *Gs* factor; for the second analysis, the CAS Successive tests were loaded on the WJ III *Gsm* factor, and so on. Based on CHC theory, one would predict that each of these integrated models would fit the data as well as the unconstrained model, that is, result in a statistically nonsignificant $\Delta\chi^2$. From the perspective of PASS theory, in contrast, one would predict that each of these modifications will result in a misspecified model and lead to a statistically significant increase in $\Delta\chi^2$.

Table 4 shows the fit associated with each of the integrated CHC models. For the first integrated model, the CAS Planning and Attention subtests were loaded on the WJ III *Gs* factor. As shown in the table, imposition of these constraints did not result in a statistically significant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 24.776$, $df = 19$, $p = .168$). Results of this model, therefore, support the CHC-derived predictions, and the hypothesis that the CAS Planning and Attention tests measure processing speed.

For the second integrated model, the Successive subtests of the CAS were loaded on the *Gsm* factor. As shown in Table 4, this model modification also resulted in a statistically nonsignificant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 11.676$, $df = 8$, $p = .166$). Results of this model, which are consistent with CHC theory and inconsistent with PASS theory, support the hypothesis that the CAS Successive tests measure short-term memory.

We then examined three additional integrated models, each of which further investigated the nature of the CAS Simultaneous tests. Results of these analyses are also shown in Table 4. For the "Integrated *Gf* and *Gv* Factors 1" model, the CAS Nonverbal Matrices subtest was loaded on the *Gf* factor and the CAS Verbal Spatial Relations and Figure Memory subtests were loaded on the *Gv* factor (cf. McGrew & Flanagan, 1998). These constraints resulted in a statistically significant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 29.671$, $df = 7$, $p < .001$), indicating that this model should be rejected. For the "Integrated *Gf* and *Gv* Factors 2" model, all the CAS Simultaneous tests were loaded on both the *Gf* and *Gv* factors. Although the $\Delta\chi^2$ for this model was not statistically significant ($\Delta\chi^2 = 6.999$, $df = 4$, $p = .136$), all of the loadings of the CAS subtests on the *Gf* factor were smaller than their loadings on the *Gv* fac-

Table 4
Comparison of the Integrated CFA Models Testing CHC Theory-Derived Hypotheses About the Nature of the CAS

Model	$\Delta\chi^2$ (*df*)	*p*	AIC
Unconstrained, nonhierarchical conjoint model	475.317 (350)[a]		705.317
Integrated *Gs* factor	24.776 (19)[b]	.168	692.093
Integrated *Gs* and *Gsm* factors	11.676 (8)[c]	.166	687.769
Integrated *Gf* and *Gv* factors 1	29.671 (7)[d]	<.001	703.440
Integrated *Gf* and *Gv* factors 2	6.999 (4)[d]	.136	686.768
Integrated *Gv* factor	10.102 (7)[d]	.183	683.871
	3.103 (3)[e]	.376	

[a]The $\Delta\chi^2$ shown for the unconstrained, nonhierarchical conjoint model is the χ^2. It is the same baseline model as in Tables 2 and 3. [b]Compared to the unconstrained nonhierarchical conjoint model. [c]Compared to the Integrated *Gs* factors model. [d]Compared to the Integrated *Gs* and *Gsm* factors model. [e]Compared to the Integrated *Gf* and *Gv* factors 2 model.

tor. For the final integrated model, the "Integrated *Gv* Factor" model, all the CAS Simultaneous tests were loaded on the *Gv* factor. As shown in Table 4, this model had an equivalent fit ($\Delta\chi^2 = 10.102$, $df = 7$, $p = .183$) to the comparison model. In addition, this model showed an equivalent fit to the "Integrated Gf and Gv factors 2" model ($\Delta\chi^2 = 3.103$, $df = 3$, $p = .376$) and the AIC for this model was smaller than that for any of the other integrated *Gf/Gv* models. As in previous analyses, these results offer partial support for either the CHC or the PASS predictions. Taken as a whole, these results further support the interpretation of the CAS Simultaneous factor as an index of visual processing, not a mixture of visual processing and fluid reasoning.

This final, integrated model represents a useful model for future research on the joint structure of the CAS and the WJ III. It is among the best fitting of the models tested (based on the AIC) and illustrates well the overlap of measurement of the two instruments.

PASS Models

This fourth series of analyses tested several models from a PASS theory perspective. For these analyses, we loaded several WJ III tests on CAS factors, based on our analyses of the task demands of these respective tests as seen from the perspective of PASS theory. In addition, we imposed several other model constraints that were consistent with PASS theory and inconsistent with CHC theory.

Attention. Table 5 shows various models using the Auditory Attention test of the WJ III. Auditory Attention is a new test that was not included in previous versions of the WJ, nor in the analyses reported above. On Auditory Attention, the examinee must point to pictures of words heard on an audiotape; response alternatives for each item consist of pictures that represent words with similar sounds (e.g., ship, zip, chip, sip). In addition, an increasing level of background noise is presented on the audiotape across items. Because of the background noise, to perform well on this test examinees must pay increasingly closer attention to the words presented on the tape. From the perspective of PASS theory, this test appears to be a classic measure of attention, which is defined as "a mental process by which the individual selectively focuses on particular stimuli while inhibiting responses to competing stimuli presented over time" (Naglieri & Das, 1997, p. 3), and requires focused, selective, and sustained attention (Naglieri, 1999a, p. 15).[6] According to PASS theory, therefore, one would predict that the Auditory Attention test should load on the Attention factor of the CAS. Unlike the CAS attention tasks, however, Auditory Attention is not speeded. In other words, if attentional skills are in fact the primary source of variation of the CAS Attention factor, and if our categorization is correct, Auditory Attention should be an integral part of this factor. In contrast, according to CHC theory, one would predict that the Auditory Attention test should load on the *Ga* factor of the WJ III. From this perspective, the central construct measured by the test is not attention, per se, but auditory processing (viz., speech/general sound discrimination or the ability to detect differences in speech sound under conditions of distraction or dis-

Table 5
Comparison of PASS versus CHC Theory-Derived Hypotheses Concerning the Nature of the Auditory Attention Test

Model	$\Delta\chi^2 (df)$	p	AIC
Auditory Attention on both Attention and *Ga*	520.860 (378)[a]		756.86
Auditory Attention on *Ga* only	1.066 (1)[b]	.302	755.926
Auditory Attention on Attention only	9 085 (1)[b]	.003	763.945

[a]The $\Delta\chi^2$ shown is the χ^2. [b]Model compared to the Auditory Attention on both Attention and *Ga* model.

tortion). Results of the three models that tested these predictions are summarized in Table 5.

The first model (Auditory Attention on both Attention and *Ga*) allowed the Auditory Attention test to load on both the CAS Attention factor and the WJ III *Ga* factor; this model represents the baseline for comparison with the next two models. For the second model, the Auditory Attention test was loaded on only the WJ III *Ga* factor. From a PASS perspective, such a constraint should result in a statistically significant degradation in model fit; from a CHC perspective, the constraint should not lead to a degradation in fit. This additional constraint did not result in a statistically significant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 1.066$, $df = 1$, $p = .302$), however. For the third model, the Auditory Attention test was loaded only on the CAS Attention factor. From a PASS perspective, such a constraint should yield a statistically nonsignificant increase in $\Delta\chi^2$ over the initial model; from a CHC perspective, this model should result in a statistically significant degradation in fit. When the Auditory Attention test was loaded only on the CAS Attention factor, however, the resulting increase in $\Delta\chi^2$ was statistically significant ($\Delta\chi^2 = 9.085$, $df=1$, $p = .003$), suggesting that this constraint should be rejected. Although these latter two models cannot be compared via the $\Delta\chi^2$ (because they are not nested), they can be compared via the AIC; as shown in the table, the AIC also supports the CHC theory prediction over the PASS theory prediction. In sum, the Auditory Attention test appears to require attention, but is nonspeeded. If so, and if the construct measured by the CAS Attention tests and factor is attention (rather than processing speed), then the Auditory Attention test should load on this factor. The fact that this test did not load on the factor provides additional evidence for the CHC-derived hypothesis that the central construct measured by the CAS "Attention" factor is not, in fact, attention.

Planning. In a similar set of analyses, we explored several models using the Planning test of the WJ III. This test requires examinees to trace a dotted figure without lifting pencil from paper and without retracing any segment. The test thus appears comparable to the CAS Planning subtests, particularly Planned Connections, because the examinee must think through each problem prior to solving it in order to answer correctly. Unlike the various tests of planning on the CAS, however, the WJ III Planning test is not speeded. From the perspective of the PASS theory, therefore, this test should measure planning without being confounded by processing speed. If the CAS Planning subtests do indeed measure planning and not processing speed, then the WJ III Planning test will load highly on the CAS Planning factor. If the WJ III Planning test primarily measures *Gv*, however, as predicted from CHC theory, it will load on the integrated CAS/WJ III *Gv* factor from Table 4 (also displayed in Figure 3).

Table 6 summarizes models that test these predictions. For the first model, the baseline model, the WJ III Planning test was

Table 6
Comparison of PASS versus CHC Theory-Derived Hypotheses Concerning the Nature of the Planning Test

Model	$\Delta\chi^2$ (*df*)	*p*	AIC
Planning test on CAS Planning factor and integrated CAS/WJ *Gv* factor	530.152 (388)[a]		746.152
Planning test on integrated CAS/WJ *Gv* factor only	1.972 (1)[b]	.160	746.124
Planning test on CAS Planning factor only	15.337 (1)[b]	<.001	759.489

[a]The $\Delta\chi^2$ shown is the χ^2. [b]Model compared to the "Planning test on CAS Planning factor and integrated CAS/WJ *Gv* factor" model.

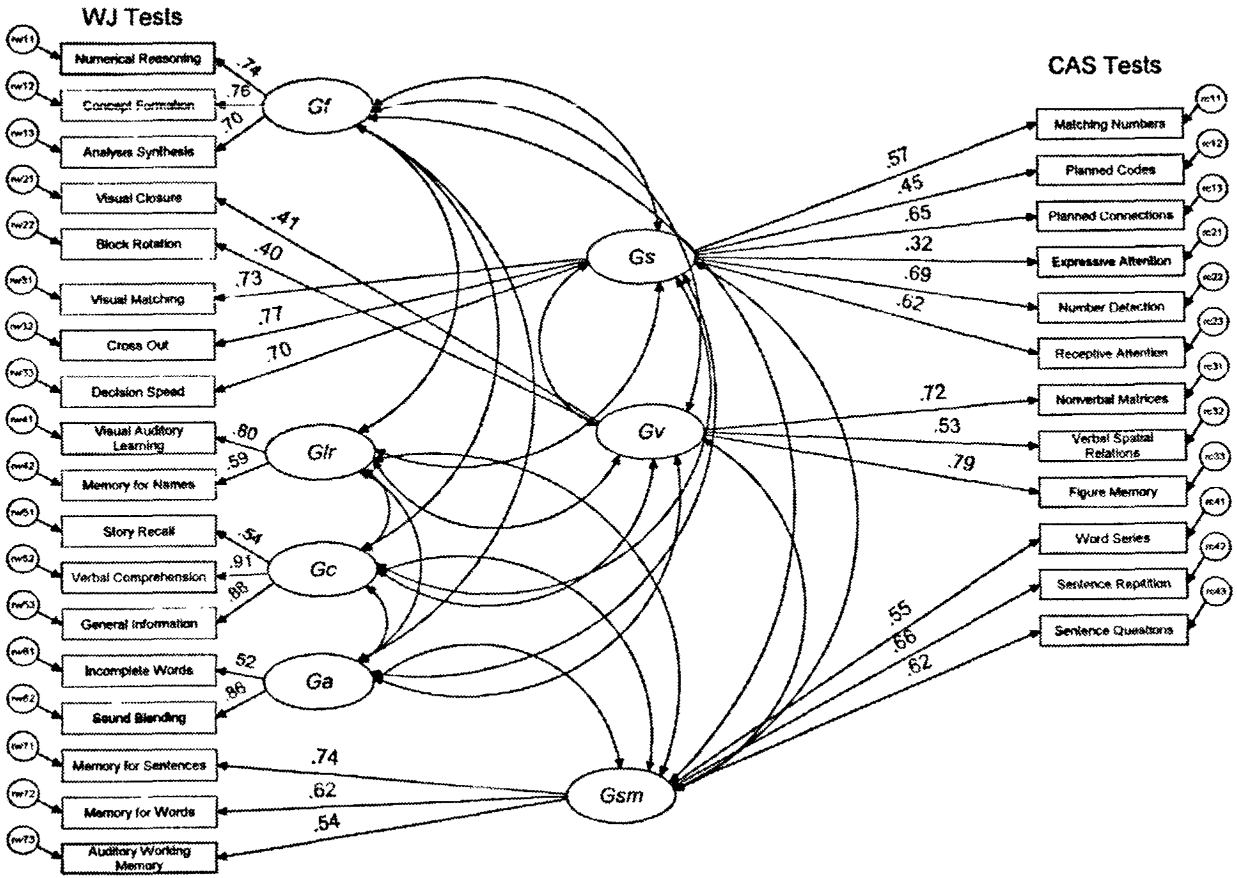

Figure 3. The final integrated CFA model of the CAS and the WJ III. The values for the factor intercorrelations are not shown to improve readability of the figure. The CAS tests fit well on the WJ III factors.

loaded on the CAS Planning factor and the integrated CAS/WJ III *Gv* factor. In the second model, the WJ III Planning test was loaded only on the integrated CAS/WJ III *Gv* factor. For the third analysis, the WJ III Planning test was loaded only on the CAS Planning factor. As can be seen in Table 6, the $\Delta\chi^2$ for the second model was not statistically significant, but the $\Delta\chi^2$ for the third model was. Again, the analyses support the CHC predictions, but do not support the PASS predictions. These findings also suggest that the WJ III Planning test does not measure the same underlying ability as the subtests on the Planning Scale of the CAS. Comparison of the AIC for the latter two models also directly supports the model in which the WJ III Planning test did not load on the CAS planning factor. To the extent that the WJ III Planning test requires planning ability, these results further indicate that the central characteristic measured by the CAS Planning tests is something other than planning ability. Results of previous analyses supported the hypothesis that the construct measured by the CAS Planning tests is not planning, but processing speed.

Simultaneous. Table 7 summarizes several models testing alternative hypotheses of the abilities measured by the WJ III Sound Blending test. For all previous analyses, the Sound Blending test was included only on the *Ga* factor. From the perspective of PASS theory, however, this test appears to be a good measure of simultaneous processing, because examinees are required to blend or synthesize word parts (phonemes or syllables) into a whole word. As Naglieri and Das (1997) stated, simultaneous processing is "a mental process by which the individual integrates separate stimuli into a single whole" (p. 4). Thus, according to PASS theory, one would predict that a model in which Sound Blending loads on the CAS Simultaneous factor should provide a significant improvement in fit over models that do not allow such loadings. From a CHC perspective, however, such a loading should not improve model fit.

Such alternative models have the additional benefit of addressing questions incompletely answered previously by this research. For the analyses reported so far, it has been difficult to differentiate *Gv* from simultaneous processing, because CHC theory and PASS theory make similar predictions. For example, after examining the model shown in Figure 3, a proponent of PASS theory might argue that the WJ III *Gv* tests and the CAS Simultaneous tests indeed measure the same thing, but what they measure is simultaneous mental processing. What is needed to make this separation is a task that measures one ability or the other (viz., *Gv* or simultaneous processing), but not both. Sound Blending appears to be such a task. Based on PASS theory, one would predict that the WJ III Sound Blending test should show a

Table 7
Comparison of PASS Theory versus CHC Theory-Derived Hypotheses About the CAS Simultaneous Scale

Model	$\Delta\chi^2$ (*df*)	*p*	AIC
Sound Blending on both WJ *Ga* and CAS Simultaneous factors	.392 (1)[a]	.531	706.925
Sound Blending only on CAS Simultaneous factor	37.929 (2)[b,c]	<.001	740.854
Sound Blending on both WJ *Ga* factor and the integrated *Gv* factor	.378 (1)[d]	.539	685.493
Sound Blending only on the integrated *Gv* factor	38.278(2)[b,c]	<.001	719.771

[a]Compared to the Unconstrained, nonhierarchical conjoint model from Table 2. [b]Compared to the previous model. [c]To estimate these models, the error variance of the Incomplete Words test was constrained to the value from the comparison model. [d]Compared to the "Integrated *Gv* factor" model from Table 4.

strong loading on the Simultaneous factor of the CAS and that the fit of the model should improve when Sound Blending is allowed to load on a Simultaneous factor. If this CAS factor does not measure simultaneous processing, however, such a model would not show a statistically significant improvement in fit and the loading of Sound Blending on Simultaneous should be inconsequential.

Table 7 shows the fit statistics for several models that tested these PASS theory-derived hypotheses. The first model allowed Sound Blending to load on both the WJ III *Ga* factor and the CAS Simultaneous factor. As can be seen in the table, allowing this additional factor loading did not result in a statistically significant improvement in fit for this model over the unconstrained, nonhierarchical conjoint model ($\Delta\chi^2 = .392$, $df = 1$, $p = .531$). In addition, the loading of Sound Blending on the Simultaneous factor was small (.12). For the second model in this series, we allowed Sound Blending to load only on the Simultaneous factor of the CAS. This model resulted in a statistically significant degradation of the fit of the model ($\Delta\chi^2 = 37.929$, $df = 2$, $p < .001$). For the third model, Sound Blending was placed on both the Ga factor and an integrated *Gv*/Simultaneous factor. This model was identical to the one shown in Figure 3, with the addition of a path from *Gv* to Sound Blending. Again, if this integrated factor represents visual processing, this model should result in no improvement in fit over the "Integrated *Gv* factor" model. If, on the other hand, this factor and the tests that load on it measure simultaneous processing, this model should show a statistically significant improvement in fit. As shown in Table 7, however, this Sound Blending on both the WJ III *Ga* factor and the Integrated *Gv* Factor model did not result in a statistically significant improvement in fit ($\Delta\chi^2 = .378$, $df = 1$, $p = .539$), and the loading of Sound Blending on *Gv* (.12) was inconsequential. Likewise, when Sound Blending was loaded only on this integrated *Gv* factor, there was a significant degradation in model fit ($\Delta\chi^2 = 38.278$, $df = 2$, $p < .001$). Hence, these analyses failed to support the PASS-theory-derived predictions, but did support the CHC-theory-derived predictions. In addition, because the Sound Blending test appears to require simultaneous processing without the necessity of visual processing, results of these analyses support the hypothesis that the CAS Simultaneous factor measures visual processing, not simultaneous mental processing.

Planning. The final analyses in this series focused on the nature of the WJ III *Gf* tests, but from the perspective of the PASS theory. According to Naglieri and Das (1997), "planning is a mental process by which the individual determines, selects, applies, and evaluates solutions to problems" (p. 2). Based on PASS theory, the WJ III *Gf* tests should be good measures of planning, because they all require the ability to form concepts and solve problems that often include novel information or procedures (cf. Kaufman & Kaufman, 1993,

Table 8
Comparison of the CFA Models Testing PASS Theory-Derived Hypotheses About the Nature of the WJ Fluid Intelligence Factor

Model	$\Delta\chi^2 (df)^a$	p	AIC
Correlation of *Gf* and Planning set to 1.00	13.971 (1)	<.001	717.288
WJ *Gf* and CAS Planning factors on a hierarchical Planning factor	37.902 (8)	<.001	727.219
WJ *Gf* and CAS Planning tests on an integrated Planning factor	57.803 (10)	<.001	743.12

[a]Models compared to the unconstrained, nonhierarchical conjoint model from Table 1.

chap. 3). To test this hypothesis, we first reanalyzed the initial, nonhierarchical conjoint model, but with the correlation between the CAS Planning factor and the WJ III *Gf* factor set to 1.00. This model tests the contention that the *Gf* and Planning factors reflect the same source of variance. As can be seen in Table 8, this constraint resulted in a statistically significant increase in $\Delta\chi^2$ over the unconstrained conjoint model ($\Delta\chi^2 = 13.971$, $df = 1$, $p < .001$), indicating that the hypothesis that *Gf* and Planning are indistinguishable should be rejected. Second, we placed both the WJ III *Gf* factor and the CAS Planning factor on a hierarchical Planning factor. This "Hierarchical Planning" model also resulted in a statistically significant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 37.902$, $df = 8$, $p < .001$). Finally, we forced the WJ III *Gf* tests to load on the CAS Planning factor rather than a WJ III *Gf* factor. Again, these changes resulted in a statistically significant increase in $\Delta\chi^2$ ($\Delta\chi^2 = 57.803$, $df = 10$, $p < .001$). Taken as a whole, results of these analyses substantiate the hypothesis that the WJ III *Gf* factor and the CAS Planning factor do not measure the same ability. If the *Gf* tests require planning ability—as would seem to be predicted by PASS theory—then these results support the hypothesis that planning is not the central construct measured by the CAS Planning tests.

Does the CAS Measure Psychometric *g*?

Kranzler and Keith (1999) and Keith and Kranzler (1999) hypothesized that a substantial psychometric *g* underlies the CAS. This hypothesis is consistent with CHC theory, which assumes that psychometric *g* underlies all cognitive tests. To examine the nature of the *g* factor on the CAS, we developed a hierarchical confirmatory model in which psychometric *g* was derived for both the CAS and the WJ III and then correlated these two factors. This model also allowed residual correlations among appropriate first-order factors (e.g., among the CAS Planning, CAS Attention, and WJ III *Gs* factors). A high correlation between the two *g* factors would support the hypothesis that the CAS measures psychometric *g*. This initial model fit well in comparison to the unconstrained conjoint model (AIC = 685.464 versus 705.317 for the unconstrained, conjoint model). Moreover, the correlation between the two *g* factors was .98. When the correlation between the two *g* factors was constrained to 1.00, the $\Delta\chi^2$ was not statistically significant ($\Delta\chi^2 = .346$, $df = 1$, $p = .556$), suggesting that the CAS *g* is statistically indistinguishable from the WJ III *g*. In addition, this model was logically equivalent to a model in which there is only one *g* factor underlying all CAS and WJ III tests. These results, therefore, support the hypothesis that the CAS measures the same psychometric *g* as does the WJ III.

Based on the relative loadings of the first-order factors on a hierarchical *g* factor, and the narrow range of abilities sampled by the CAS, Kranzler and Keith (1999) also hypothesized that "the FS score on the CAS may not be the best estimate of psychometric *g*....The best estimate of *g* on the CAS appears to be the Simultaneous Scale" (p. 139). It is important to note that this hypothesis was derived from their analyses, not from CHC theory, per se. To test this hypothesis, we compared the loading of the CAS FS score on a psychometric *g* factor in one model with the loading of the CAS Simultaneous Scale score in another model. For both models, the WJ III first-order factors were also loaded on the hierarchical *g* factor. Results of these analyses indicated that the CAS FS score and Simultaneous Scale score had similar loadings on psychometric *g*. The CAS FS loaded .79 versus .77 for the Simultaneous Scale. The FS and Simultaneous Scale scores of the CAS, therefore, appear to be equivalent measures of psychometric *g*. Additional research will be needed to determine whether or not they provide "good" measures of *g* (see Jensen & Weng, 1994).

Discussion

Considerable disagreement has surrounded the factors measured by preliminary batteries of PASS tests and the CAS (e.g., Carroll, 1995; Kranzler & Keith, 1999; Kranzler & Weng, 1995a, 1995b; Naglieri et al., 1991). Kranzler and Keith's (1999) recent study of the standardization data suggested

Table 9
Summary of Findings: Support for Cattell-Horn-Carroll versus PASS Theory-Derived Hypotheses about the Nature of the CAS

Model Series and Research Questions	CHC Theory Supported	PASS Theory Supported
Conjoint Models		
Planning = Planning or Processing Speed?	Processing Speed	
Attention = Attention or Processing Speed?	Processing Speed	
Successive = Successive or Short-Term Memory?	Short-Term Memory	
Simultaneous = Simultaneous or Fluid Intelligence & Visual Processing?	?	?
Hierarchical Conjoint Models		
Planning = Planning or Processing Speed?	Processing Speed	
Attention = Attention or Processing Speed?	Processing Speed	
Successive = Successive or Short-Term Memory?	Short-Term Memory	
Simultaneous = Simultaneous or Fluid Intelligence & Visual Processing?	Yes – Visual Processing No – Fluid Intelligence	?
Integrated CHC Models		
Planning = Planning or Processing Speed?	Processing Speed	
Attention = Attention or Processing Speed?	Processing Speed	
Successive = Successive or Short-Term Memory?	Short-Term Memory	
Simultaneous = Simultaneous or Fluid Intelligence & Visual Processing?	Yes – Visual Processing No – Fluid Intelligence	?
Integrated PASS Models		
Planning = Planning or Processing Speed?	Processing Speed	
Attention = Attention or Processing Speed?	Processing Speed	
Simultaneous = Visual Processing (these models also clarify all of the cells containing question marks)	Visual Processing	
Psychometric *g* Models		
Does the CAS Measure Psychometric *g*?	Yes	

Note. Entries that appear in the CHC column show support and the nature of the support for CHC theory; those that appear in the PASS column show support for PASS theory.

that the CAS lacks structural fidelity, indicating that the CAS does not measure what Naglieri and Das (1997) intended it to measure. Instead, Kranzler and Keith (1999) hypothesized that the CAS FS and Scales do not measure PASS processes, but are better understood within the framework of CHC theory as measures of psychometric *g*, processing speed, short-term memory span, and a mixture of fluid intelligence/broad visualization. The aim of this study was to evaluate these hypotheses by testing a series of competing predictions about the nature of the CAS, with those predictions drawn from PASS theory and from CHC theory. To do so, we conducted joint CFAs of the CAS and a measure of intelligence designed to measure the broad cognitive abilities in CHC theory—the WJ III (Woodcock et al., 2001).

Several sets of joint CFAs were conducted to test Kranzler and Keith's (1999) hypotheses. In the first set of analyses, we conducted a series of first-order conjoint CFAs in which the factors on the CAS and WJ III were initially allowed to correlate freely. Various constraints were then imposed on the models to test specific hypotheses based on predictions from the CHC or PASS theoretical perspectives. In the second set of analyses, we examined models in which the first-order factors of the CAS and WJ III loaded on higher-order factors in CHC theory (e.g., *Gs*, *Gsm*). In the third set, we conducted a series of integrated CFAs in which we examined the viability of loading CAS subtests on WJ III factors based on CHC theory. For the fourth set of analyses, we reversed this process and tested a series of models derived from PASS theory, including the addition of several tests that should behave differently depending on whether the CAS is measuring the PASS processes or the broad cognitive abilities in CHC theory. In the fifth and final set of analyses, we examined the relation between the psychometric *g* factor on each respective test. Results of these joint CFAs are summarized in Table 9. As shown in this table, our findings consistently supported the predictions from CHC theory, and consistently supported Kranzler and Keith's (1999) hypotheses about the constructs measured by the CAS. These findings also provided independent support for Flanagan and colleagues' classifications of the CHC abilities underlying the CAS tests (Flanagan et al., 2000; McGrew & Flanagan, 1998). Likewise, the findings consistently refuted hypotheses derived from PASS theory.

Do the Planning and Attention Scales Measure Processing Speed (*Gs*)?

Yes. In the conjoint CFAs, when the factors of the CAS and WJ III were allowed to freely correlate, the correlations between Planning and Attention and *Gs* approached unity. Specifically, the correlation between Planning and *Gs* was .98 and between Attention and *Gs* was .88. Constraining the correlations between Planning and Attention and *Gs* to unity resulted in an equivalent model fit in comparison to the unconstrained model. In contrast, constraining the correlation between these factors to the average interfactor correlation resulted in a significant degradation in model fit. We also examined a model in which the CAS Planning and Attention factors and the WJ III *Gs* factor were loaded on a higher-order *Gs* factor. The fit of this model was equivalent to that for the unconstrained model. Results of the integrated CFAs were similar. Loading the CAS Planning and Attention subtests on the WJ III *Gs* factor did not result in a significant degradation in model fit. Thus, the construct underlying the Planning and Attention factors of the CAS appears to be broad processing speed (*Gs*) as defined by CHC theory.

This hypothesis was further supported by testing a model that added a new test that appears to measure attention as defined in the PASS theory, without requiring processing speed. We examined the relative fit of models in which Auditory Attention, a new test on the WJ III that was not included in other analyses in this research, was loaded on the CAS Attention factor and the WJ III Auditory Processing (*Ga*) factor. Auditory Attention appears to meet the definition of a measure of attention processes in the PASS theory, but without the heavy emphasis on speed required by the CAS Attention subtests (see Naglieri

& Das, 1997, pp. 3-4). We hypothesized that if the CAS Attention factor truly measures attention as Naglieri and Das intended and not processing speed, then confining Auditory Attention on this factor should not result in a degradation in model fit. When we conducted these CFAs, a statistically significant degradation in model fit was observed when the Auditory Attention test was loaded on the CAS Attention factor, but not when it was loaded on the WJ III *Ga* factor.

Similar results were obtained when we examined the fit of models in which the new Planning test of the WJ III was loaded on the CAS Planning factor (consistent with PASS theory) and the integrated *Gv* factor (consistent with CHC theory). Although this new test is not as thoroughly understood as the other WJ III tests, results of this set of analyses are wholly consistent with the other CFA results: CHC theory was again supported; PASS theory was not. Specifically, the WJ III Planning test—quite similar to the CAS Planned Connections subtest—did not load on the Planning factor as predicted by PASS theory, most likely because it is not highly speeded like the CAS planning tests.

In sum, results of this set of joint CFAs repeatedly and consistently substantiate the hypothesis that the CAS Planning and Attention factors should be interpreted as processing speed. Not only are correlations between these CAS factors and the WJ III processing speed factor near unity, but nonspeeded tests apparently requiring attention and planning on the WJ III do not measure the same constructs as the CAS Attention and Planning factors. These results contradict predictions derived from PASS theory, but bolster those based on CHC theory.

Does the Successive Scale Measure Short-Term Memory (*Gsm*)?

Yes. When the CAS and WJ III factors were allowed to correlate freely, the correlation between the CAS Successive factor and the WJ III *Gsm* factor was not significantly different from unity. When we constrained the correlation between these two factors to 1.00, this model provided an equivalent fit to the unconstrained model. Constraining this interfactor correlation to a moderate level, as predicted by PASS theory (see Das et al., 1994, p. 126), resulted in a significant degradation in model fit. We also examined models in which the CAS Successive factor and the WJ III *Gsm* factor were loaded on a higher-order *Gsm* factor. The fit of this model was equivalent to that for the unconstrained model. Moreover, the fit provided by a model in which the CAS Successive subtests were integrated with the *Gsm* factor also resulted in an equivalent fitting model.

Taken as a whole, results of this set of joint CFAs consistently demonstrate that the CAS Successive factor is indistinguishable from the WJ III *Gsm* factor, and that the WJ III *Gsm* tests and the CAS Successive tests form a coherent *Gsm* factor. The correlation of 1.00 between these two factors, their loading on a hierarchical *Gsm* factor, and the good fit of the integrated *Gsm* factor all indicate that the CAS Successive tests and the WJ III *Gsm* tests reflect the same underlying source of variance. Once again, the close relation between the Successive and *Gsm* factors is consistent with predictions based on the CHC theory and refutes those based on the PASS theory.

Does the Simultaneous Scale Measure a Mixture of Fluid Intelligence (*Gf*)/Broad Visualization (*Gv*)?

No. The Simultaneous Scale appears to measure broad visualization (*Gv*). When discussing results of our joint CFAs of the Simultaneous factor, it is important to note that, based on CHC theory, Kranzler and Keith (1999) predicted that the Simultaneous factor on the CAS would correlate moderately with the *Gv* and *Gf* factors on the WJ III, but that, based on PASS theory, Das et al. (1994) suggested that the Simultaneous factor would be related to *Gv* as well. To the best of our knowledge, Naglieri, Das, and their colleagues have not addressed the possible relation between the Simultaneous and *Gf* factors. Predictions based on CHC and PASS theory, therefore, overlap and cannot always be interpreted as clear support for one theory over the other.

In the unconstrained model, the Simultaneous factor correlated .68 with *Gv* and .77 with *Gf*. We also examined the fit provided by a set of three hierarchical conjoint CFAs. The best-fitting of these models suggested that the CAS Simultaneous factor and the WJ III *Gv* factor were reflections of the same broad ability. The relation between Simultaneous and *Gv* factors was also supported in a set of three integrated models. In these models, loading all the subtests from the Simultaneous Scale of the CAS onto a *Gv* factor provided an equivalent fit to the data as did the comparison model. Based on this set of CFAs, these findings could be interpreted either as support for Kranzler and Keith's CHC theory predictions or as support for PASS theory predictions.

Results of several CFAs based on predictions from the PASS perspective shed light on this ambiguity, however. For example, the Sound Blending test of the WJ III, a test that meets the definition of a measure of simultaneous mental processing, did not load on a CAS Simultaneous factor. Likewise, the WJ III Planning test, a test requiring both visual processing and planning, but not simultaneous mental processing, loaded on an integrated *Gv* factor that included both the WJ III *Gv* and the CAS Simultaneous tests. If our categorizations are correct, results of these analyses, therefore, support the CHC-derived hypothesis that the CAS Simultaneous tests (and the WJ III *Gv* tests) measure visual processing, not simultaneous mental processing. Therefore, it follows that the CAS Simultaneous tests are most appropriately interpreted as measures of broad visualization (*Gv*).

Does the CAS FS Score Measure Psychometric *g*?

Yes. For any particular battery of mental tests, the *g* factor is estimated equally well by the first unrotated principal factor, the first unrotated principal component, the single highest-order factor in a Schmid-Leiman hierarchical factor analysis, and CFA methods of factor analysis (see Jensen, 1998; Jensen & Weng, 1994; Ree & Earles, 1991). In CHC theory, *g* is the most general intellectual ability, and subsumes all other mental abilities. In contrast, *g* has no place in PASS theory, and the CAS was not designed to assess psychometric *g* (Naglieri, 1999a, 1999b).

To examine whether the CAS measures the same psychometric g as the WJ III, we developed a model with a hierarchical *g* for both the CAS and the WJ III. These *g* factors were then correlated. The correlation between these two factors was .98. Constraining the correlation between the two *g* factors to 1.00 resulted in an equivalent fit. Results of our analyses, therefore, suggest that the *g* factor underlying the CAS is, for all intents and purposes, indistinguishable from the *g* factor underlying the WJ III. The CAS appears to require the same "general intelligence" as do all other cognitive test batteries.

We also tested Kranzler and Keith's (1999) hypothesis that the Simultaneous Scale score, not the FS score, is the best estimate of *g* on the CAS. To do so, we compared the loading of the CAS FS score and the CAS Simultaneous Scale score on the *g* factor in two separate models. Results of these analyses indicated that the FS score loaded .79 on *g*, whereas the Simultaneous Scale score loaded .77. The slightly higher loading of the FS score on *g* suggests that the Simultaneous Scale is not a better estimate of *g* than is the CAS overall score.

Study Limitations and Future Research

Sample. One potential limitation of this research is the sample size. Readers familiar with rules of thumb for exploratory factor analysis (e.g., 10 participants per measured variable; although see Goodwin & Goodwin, 1999, for a critique of this "rule") will wonder whether our sample size of 155 is adequate to test the hypotheses of interest in this research. In the absence of contrary data, many researchers have adopted similar rules in confirmatory factor analysis and structural equation modeling (SEM). The primary concern with small samples is the ability to differentiate good from poor models; with small sample sizes, one may not have adequate power to reject inadequate models, or to differentiate good from bad models. More recently, however, research has shown that power in CFA and SEM is not only dependent on sample size, but on the number of indicators (more variables are better) and the constraints imposed (more constraints or *df* are better) (Loehlin, 1998; MacCallum, Browne, & Sugawara, 1996; Marsh, Hau, Balla, & Grayson, 1998). In fact, it is possible to calculate the power (the probability of correctly rejecting a false null hypothesis) of a CFA/SEM model from the sample size and the degrees of freedom. The highly constrained models ($df \geq 350$) used in this research had ample power. For example, the model shown in Figure 1 (among our least constrained, least powerful, models) had an estimated power of .999 (alpha = .01) to reject the null hypothesis of a close fit of the model to the data (MacCallum et al., 1996). Thus, our sample size did not result in too little power. Additionally, the rejection of many of the hypotheses tested here illustrates that our model comparisons likewise had ample power.

Another possible problem with small samples is that the models may be unstable, or produce inaccurate parameter estimates. Our sample size is within the range recommended by methodologists, however (Loehlin, 1998). Furthermore, the consistency of estimates (e.g., factor loadings) across a variety of models argues for the stability of our models and findings.

The sample characteristics are another potential limitation of this research. Because participants in this research consisted of a diverse group of general education students in terms of race/ethnicity, age, and sex, critics might argue that the proportion of students in our sample in each of these categories differed from the standardization samples of the CAS and the WJ III. Perhaps, one might argue, our findings are simply an artifact of the sample. Research on test bias for over 25 years, however, has clearly shown that the constructs measured by intelligence tests are the same across all groups of English-speaking children born and raised in the United States (e.g., Reynolds, Lowe, & Saenz, 1999). Moreover, at the current time there are no data to suggest that either the CAS or WJ III is biased (i.e., measures different constructs) across groups. Thus, the demographic characteristics of our sample should not limit the generalizability of our results on the constructs measured by the CAS.

These latter two potential limitations (stability and sample characteristics) can be evaluated empirically, as well. To do so, we constrained the (unstandardized) factor loadings for each test to the values obtained in CFAs of each test's standardization data. For the CAS, we used the factor loadings for ages 8-10 (from Kranzler & Keith, 1999); for the WJ III we used the loadings for ages 8-11.[7] The model with these constraints was compared to the unconstrained model (Figure 1). The $\Delta\chi^2$ was not statistically significant ($\Delta\chi^2 = 29.722$, $df = 19$, $p = .055$), meaning that the factor structures obtained for the CAS and the WJ III in the present research were statistically indistinguishable from those that would be obtained with each test's standardization data. These findings show both that the present results are stable and that they are not idiosyncratic to these data. In sum, we have shown that our analyses have adequate power to test the hypotheses of interest, that the results are stable, and that the results are likely similar to those that would be obtained with other larger and more representative samples.

This research could also be criticized because we used a newer version—the upcoming third edition—of the WJ Tests of Cognitive Ability to measure the CHC abilities of interest. We reiterate that the purpose of this research was not to test the validity of the WJ III; instead, the WJ III tests were used to provide markers for the CHC abilities. There is ample evidence that the previous edition of the WJ (the WJ-R) measures the broad constructs of CHC theory, and does so across a wide age span (e.g., Bickley et al., 1995). Of course, many of the tests used in the present research are the same as those from the WJ-R. Furthermore, because the WJ III was designed to measure the broad CHC abilities more completely through two or three qualitatively different narrow ability indicators of the respective broad abilities, it should provide a better operationalization of the CHC theory than the WJ-R. Finally, CFA of the WJ III suggests that this latest version of the test indeed provides valid measures of the CHC constructs (McGrew & Woodcock, 2001).

Finally, critics could argue that the predictions we derived from PASS theory are not accurate. For example, despite our contention that Sound Blending should be considered a Simultaneous task from the viewpoint of PASS theory, proponents of PASS theory may disagree. Nonetheless, all of our PASS-theory-derived predictions were based on the writing of Naglieri, Das, and their colleagues, the primary proponents of PASS theory and the authors of the CAS. Likewise, all categorizations of tests as measuring PASS constructs were based on definitions of those constructs by the CAS authors. If those categorizations or predictions are incorrect, then the definitions and explications of PASS theory must also be questioned. Idiosyncratic definitions of constructs that do not allow for generalization to new tasks do not fulfill the requirements of valid theory. Finally, all of our joint CFAs in this research were based on a priori predictions from several sources (e.g., Carroll, 1995; Das et al., 1994; Keith & Kranzler, 1999; Kranzler & Keith, 1999; McGrew & Flanagan, 1998; Naglieri, 1999a, 1999b; Naglieri & Das, 1997). Post hoc explanations of these results should carry very little weight until verified by independent, empirical data.

Further research should be conducted concerning the nature of the CAS Successive tests. This research consistently supported CHC-derived predictions concerning the nature of the CAS Successive tests, but not those derived from PASS theory. These findings thus support the CHC interpretation of these tests as measures of short-term memory rather than successive mental processing. For all other areas, however, it was possible to go beyond testing such predictions via the analysis of tests that measured one skill but not the other. Additional support was garnered for the CHC interpretation of the CAS Planning tests (as measures of processing speed) via the analysis of a test that required planning but not speed (the WJ Planning test). Likewise, we analyzed a test that required simultaneous processing but not visual processing and one that required attention but not processing speed, with all such analyses supporting the CHC interpretation of CAS tests. All CAS Successive tests require memory skills, however, and it could be argued that the WJ III tests used here require successive processing. Future research should analyze the CAS Successive tests in conjunction with one or more successive tests unconfounded with memory skills, or with memory tests unconfounded with successive skills.

Finally, it is important to review the strength of the evidence presented here. We were able to reject nearly all of the predictions generated from PASS theory, and all of the firm predictions from that theory. For that reason, and given that our predictions were true to the theory, we can state with a high degree of confidence that the CAS does not measure the PASS abilities. In contrast, we were able to reject few of the predictions generated from CHC theory, thus supporting the CHC interpretation of the CAS. We cannot state with the same degree of assurance, however, that the CAS definitely measures the CHC abilities. There may be other theoretical orientations, not tested here, that explain the relations among these tests as well as, or better than, CHC theory. CHC theory certainly does provide a better orientation for explaining CAS performance than does PASS theory, and at the present time appears to be the best such theoretical orientation for explaining CAS performance.

Conclusion

Taken as a whole, results of the joint CFAs conducted in this study fail to substantiate the construct validity of the CAS as a measure of PASS processes. Indeed, results of this study repeatedly and consistently support virtually all of Kranzler and Keith's (1999) and McGrew and Flanagan's (1998) counterinterpretations of the constructs measured by the CAS from the perspective of the CHC theory. Results of this study support the alternative explanation of the factor structure of the CAS in which the Planning/Attention Scales are combined and interpreted as measures of processing speed (cf. Carroll, 1995; Keith & Kranzler, 1999; Kranzler & Keith, 1999; Kranzler et al., 2000; Kranzler & Weng, 1995a). Our findings also support the interpretation of the Successive Scale as a measure of short-term memory span and the Simultaneous Scale as a measure of broad visualization. Finally, although the FS score was intended as a mere practical device to satisfy state regulations (see Naglieri, 1999b, p. 147), results of our study indicate that the CAS measures the same psychometric *g* as do other tests of intelligence. Overall, results of this study suggest that the CAS is neither broader in scope than other traditional intelligence tests, nor a ground-breaking measure of unique psychological constructs, as its authors maintain (Naglieri, 1999a). On the contrary, the CAS appears to measure adequately psychometric *g* and three basic cognitive abilities, which is quite comparable to most other contemporary IQ tests. Moreover, the cognitive abilities that appear to be measured by the CAS (viz., psychometric *g*, processing speed, short-term memory span, and broad visualization) have long been recognized in research on the structure of human intelligence (for a review, see Carroll, 1993).

In addition to these empirical results, this study also demonstrates the utility of the CHC theory as a framework for understanding the cognitive constructs measured by new tests of intelligence, such as the CAS. The alternative explanation of the structure of the CAS is supported by CHC theory (Flanagan et al., 2000; McGrew & Flanagan, 1998) and strong empirical evidence (Carroll, 1995; Keith & Kranzler, 1999; Kranzler et al., in press; Kranzler & Keith, 1999; Kranzler & Weng, 1995a). We therefore urge practitioners to use the CHC theory to interpret the scaled scores on the CAS, not PASS theory. From the perspective of CHC theory, the FS score on the CAS appears to reflect general intelligence, although it is untested how well the FS score (as opposed to a *g* factor) measures *g* compared to other tests. This implies that the CAS FS score may be useful when determining eligibility for special education and related services (e.g., learning disabilities, mental handicaps, intellectual giftedness). Nonetheless, because the CAS does not appear to be a valid measure of the four PASS processes, use of the CAS for identifying learning strengths and weaknesses, making differential diagnoses, and designing or implementing treatment, instructional, and remedial programs based on ipsative analysis of CAS Scale scores according to PASS theory is not warranted. Such scores may be more validly interpreted from CHC theory, however. The difference is important in that psychologists who use the CAS for developing programs are likely to provide different programs if a weakness on, for example, the Planning scale represents difficulties in processing speed rather than planning ability. Finally, given that the CAS is best understood within the context of CHC theory, practitioners who use the CAS should consider using the principles of the CHC Cross-Battery approach to supplement CAS-based assessments (Flanagan et al., 2000; Flanagan & Ortiz, 2000, McGrew & Flanagan, 1998). The final, integrated model shown in Figure 3, for example, is quite consistent with this approach.

Strong programs of construct validation consist of three stages: substantive, structural, and external (see Benson, 1998). The substantive stage concerns the operationalization of constructs specified by a psychological theory in behavioral terms (i.e., observable variables). The structural stage aims "to determine the extent to which the observed variables covary among themselves, and how they covary with the intended structure of the theoretical domain" (Benson, 1998, p. 13). In the last stage, the external stage, the meaning of scores on observable variables is substantiated by examining the degree to which the pattern of relations between test scores and external criteria is both rational and consistent with construct theory (e.g., Messick, 1995). Establishing the construct validity of any instrument requires a range of evidence gathered across all three stages and, ideally, with multiple methods within each stage. As Benson (1998) stated, results of research at "each stage either leads to the next in building evidence for the construct validity interpretation of test scores or suggests the previous stage *should be reevaluated*" (p. 15, emphasis added).

Results of research on the construct validity of the CAS by independent researchers at the structural stage, however, have consistently failed to support the construct validity of the CAS and preliminary batteries of PASS tasks (Carroll, 1995; Keith & Kranzler, 1999; Kranzler & Keith, 1999; Kranzler et al., in press; Kranzler & Weng, 1995a, 1995b). Kranzler and Keith's (1999) reanalyses of the CAS standardization data revealed that the CAS lacks structural fidelity, thereby indicating that the CAS does not measure the PASS processes. Their substantive conclusions were recently replicated by Kranzler et al. (in press). Results of the present study provide evidence refuting the construct validity of the CAS at the external stage. The pattern of relations between scores on the CAS and WJ III was neither rational—from the perspective of PASS theory—nor consistent with construct theory underlying the CAS. Results of the present study, however, were predicted by, and are easily explained by, CHC theory.

Results of independent research on the construct validity of the CAS, therefore, underscore the need for further work at the substantive stage. At this stage, every psychological construct is seen to reflect two domains—a theoretical domain and an empirical domain (Benson, 1998). The theoretical domain represents all that is known about the construct (e.g., PASS theory), whereas the empirical domain involves its operationalization in behavioral terms (e.g., FS, PASS Scale, and subtest scaled scores). Results of independent research conducted thus far on the construct validity of the CAS leads to one of the following three conclusions: Either the CAS is based on an invalid theory that is at odds with nature (i.e., what is known about individual differences in human cognitive ability), or the tests developed to assess the PASS processes are invalid operationalizations of those processes, or both. Although the factor analytic data presented in the present study may not directly address the verisimilitude of a neuropsychological theory of brain function, it is clear that the PASS theory not only fails to explain the relations among the scores on the CAS, but also its relations with another test of intelligence, the WJ III. Hence, the PASS theory cannot account for what is known about the structure of cognitive abilities that has been accumulated over almost 100 years of research (see Carroll, 1993). Based on the available empirical evidence, therefore, the PASS model appears untenable as a theory of individual differences in intelligence.

References

American Psychological Association. (1992). *Ethical principles of psychologists and code of conduct.* Washington, DC: Author.

Arbuckle, J. L., & Wothke, W. (1999). *Amos 4.0 user's guide.* Chicago: Smallwaters.

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practices, 17,* 10-17.

Bickley, P. G., Keith, T. Z., Wolfle, L. M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intelligence across the life span. *Intelligence, 20,* 309-328.

Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling, 7,* 461-483.

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

Carroll, J. B. (1995). [Review of *Assessment of cognitive processes: The PASS theory of intelligence*]. *Journal of Psychoeducational Assessment, 13,* 397-409.

Carroll, J. B. (1997). The three-stratum theory of cognitive abilities. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 122-130). New York: Guilford.

Daniel, M. H. (1997). Intelligence testing: Status and trends. *American Psychologist, 52,* 1038-1045.

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

Flanagan, D P., Genshaft, J. L., & Harrison, P. L. (1997). *Contemporary intellectual assessment: Theories, tests, and issues.* New York: Guilford.

Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Wechsler Intelligence Scales and Gf-Gc theory: A contemporary approach to interpretation.* Boston: Allyn & Bacon.

Flanagan, D. P., & Ortiz, S. O. (2000) *Essentials of cross-battery assessment.* New York: Wiley.

Goodwin, L. D., & Goodwin, W. L. (1999). Measurement myths and misconceptions. *School Psychology Quarterly, 14,* 408-427.

Horn, J L (1994) Theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), *Encyclopedia of intelligence* (pp. 443-451). New York. Macmillan.

Jensen, A. R (1998). *The g factor.* Westport, CT. Praeger.

Jensen, A. R., & Sinha, S. N. (1993). Physical correlates of human intelligence. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence* (pp. 139-242). Norwood, NJ: Ablex.

Jensen, A. R., & Weng, L. (1994). What is a good g? *Intelligence, 18*, 231-258.
Joseph, L. M. (1999). An interview with Dr. Jack A. Naglieri: Intelligence, intelligence testing and the Cognitive Assessment System (CAS) *NASP Communiqué, 28*(2), 8-10.
Kaufman, A. S., & Kaufman, N. L (1993). *Manual: Kaufman Adolescent and Adult Intelligence Test.* Circle Pines, MN: American Guidance Service.
Keith, T. Z. (1997). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan, J L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 373-402). New York: Guilford.
Keith, T. Z., & Kranzler, J. H (1999) The absence of structural fidelity precludes construct validity Rejoinder to Naglieri on what the Cognitive Assessment System does and does not measure. *School Psychology Review, 28*, 303-321.
Kranzler, J. H., & Keith, T. Z. (1999) Independent confirmatory factor analysis of the Cognitive Assessment System (CAS): What does the CAS measure? *School Psychology Review, 28*, 117-144.
Kranzler, J. H., Keith, T. Z., & Flanagan, D P. (2000). Independent examination of the factor structure of the Cognitive Assessment System (CAS): Further evidence challenging the construct validity of the CAS. *Journal of Psychoeducational Assessment, 18*, 143-159.
Kranzler, J. H., & Weng, L. (1995a) The factor structure of the PASS cognitive tasks: A reexamination of Naglieri et al. *Journal of School Psychology, 33*, 143-157.
Kranzler, J. H., & Weng, L. (1995b) A reply to the commentary by Naglieri and Das on the factor structure of a battery of PASS cognitive tasks. *Journal of School Psychology, 33*, 169-176.
Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis* (3rd ed.). Hillsdale, NJ: Erlbaum.
MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Science, 1*, 130-149.
Marsh, H. W., Hau, K-T, Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research, 33*, 181-220.
McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive G_f-G_C framework. In D. P. Flanagan, J L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp 151-179). New York: Guilford.
McGrew, K. S. & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR): Gf-Gc cross-battery assessment.* Boston: Allyn & Bacon.
McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual: Woodcock-Johnson III.* Itasca, IL: Riverside
Messick, S. (1995) Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning *American Psychologist, 50*, 741-749.
Naglieri, J A (1997) Planning, attention, simultaneous, and successive theory and the Cognitive Assessment System: A new theory-based measure of intelligence. In D. P Flanagan, J L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment. Theories, tests, and issues* (pp. 247-267). New York: Guilford
Naglieri, J A (1999a) *Essentials of CAS assessment.* New York: Wiley.
Naglieri, J A (1999b) How valid is the PASS Theory and CAS? *School Psychology Review, 28*, 145-162.
Naglieri, J. A, Braden, J. & Gottling, S (1993). Confirmatory factor analysis of the planning, attention, simultaneous, successive (PASS) cognitive processing model for a kindergarten sample. *Journal of Psychoeducational Assessment, 11*, 259-269.
Naglieri, J A, & Das, J P (1997) *Cognitive assessment system Interpretive handbook* Itasca, IL. Riverside.
Naglieri, J. A, Das, J P, & Jarman, R. F. (1990). Planning, attention, simultaneous, and successive cognitive processes as a model for assessment *School Psychology Review, 19*, 423-442.
Naglieri, J A, Das, J P, Stevens, J J., & Ledbetter, M F. (1991) Confirmatory factor analysis of planning, attention, simultaneous, and successive cognitive processing tasks *Journal of School Psychology, 29*, 1-17.
Ree, M., & Earles, J A. (1991) The stability of convergent estimates of g *Intelligence, 15*, 86-89.
Reynolds, C R., Lowe, P. A, & Saenz, A L. (1999). The problem of bias in psychological assessment. In C. R. Reynolds & T. B Gutkin (Eds), *The handbook of school psychology* (3rd ed, pp. 549-595). New York: Wiley.
Woodcock, R W., & Johnson, M. B (1989). *Woodcock-Johnson Psycho-Educational Battery–Revised: Tests of Cognitive Ability.* Chicago: Riverside.
Woodcock, R. W, McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities.* Itasca, IL. Riverside.

Footnotes

[1]Although Naglieri, Das, and their colleagues also have reported the results of CFAs of preliminary batteries of PASS tests (e g., Naglieri, Braden, & Gottling, 1993; Naglieri et al., 1991), Naglieri (1999b) argued that the results of these studies are irrelevant to the CAS because they are based on different tests (for further discussion, see Keith & Kranzler, 1999, p. 146)

[2]Naglieri did not respond to invitations by Keith and Kranzler (1999) in their article (p 319) and on a public school psychology listserve to submit predictions concerning the results of such analyses He also declined a private invitation by the first author of the present study to make such predictions

[3]Using the $\Delta\chi^2$ criterion, the more parsimonious of two nested models was accepted if the change in chi-squared was not statistically significant ($p > .05$); if the more parsimonious model resulted in a statistically significant increase in $\Delta\chi^2$ ($p < .05$), the less parsimonious model was accepted The AIC was used to compare nonnested models, smaller AIC values suggest superior models.

This initial, unconstrained, model provided an adequate fit to the data according to the stand-alone fit indices shown in Table 2 (e.g., RMSEA = .048); we could not reject the null hypothesis of a close fit of the model to the data (MacCallum, Brown, & Sugawara, 1996). Given the complexity and the highly constrained nature of our models, in which subtests from the two batteries were not allowed to cross-load on the same factors, we were surprised at how well these models fit according to these stand-alone indices.

[4]The actual point value of the correlation was greater than 1.00, but was not statistically significantly different from 1.00; the correlation should thus be considered a maximum of 1.00.

[5]Although several factor loadings were greater than 1.00, they were not statistically significantly different from 1.00.

[6]Reliance on an analysis of the task demands of a test may appear subjective at first glance, but it is a requirement of all factor analysis.

[7]We are grateful to Dr. Kevin McGrew for providing these data.

Timothy Z. Keith, PhD, is the Arthur L. and Lea R. Powell Professor of Psychology and Schooling at Alfred University. His research focuses on understanding the influences on school learning and the nature and measurement of intelligence. Tim also has methodological interests in structural equation modeling and confirmatory factor analysis.

John H. Kranzler, Ph.D., is Professor of Educational Psychology and Director of the School Psychology Program at the University of Florida. His primary research interests include the nature, development, and assessment of human cognitive abilities.

Dawn P. Flanagan, Ph.D., is Associate Professor in the Department of Psychology at St. John's University in New York. Her research interests include understanding the nature and structure of intelligence, psychoeducational and preschool assessment, and professional issues in school psychology.

Appendix

Factor Intercorrelations for the Unconstrained Conjoint Model

	Planning	Successive	Simultaneous	Attention	*Gsm*	*Ga*	*Gc*	*Glr*	*Gs*	*Gv*	*Gf*
Planning	1.00										
Successive	0.45	1.00									
Simultaneous	0.66	0.33	1.00								
Attention	0.91	0.27	0.62	1.00							
Gsm	0.43	1.06*	0.36	0.21	1.00						
Ga	0.37	0.52	0.57	0.34	0.64	1.00					
Gc	0.64	0.53	0.75	0.50	0.60	0.65	1.00				
Glr	0.54	0.23	0.64	0.38	0.46	0.41	0.57	1.00			
Gs	0.98	0.36	0.67	0.88	0.36	0.33	0.51	0.38	1.00		
Gv	0.30	0.26	0.68	0.32	0.19	0.20	0.35	0.48	0.45	1.00	
Gf	0.68	0.48	0.77	0.44	0.62	0.58	0.78	0.70	0.63	0.41	1.00

*The value reported was 1.06, but is not statistically significantly different from 1.0.