Psychology in the Schools, Vol. 44(2), 2007
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/pits.20211

JOINT CONFIRMATORY FACTOR ANALYSIS OF THE DIFFERENTIAL ABILITY SCALES AND THE WOODCOCK-JOHNSON TESTS OF COGNITIVE ABILITIES–THIRD EDITION

SARAH SANDERS AND DAVID E. MCINTOSH

Ball State University

MARDIS DUNHAM

Murray State University

BARBARA A. ROTHLISBERG AND HOLMES FINCH

Ball State University

This study examined the underlying constructs measured by the Differential Ability Scales (DAS; C.D. Elliott, 1990a) as they relate to the Cattell-Horn-Carroll (CHC) Theory (K.S. McGrew, 1997) of cognitive abilities. The DAS and Woodcock-Johnson Tests of Cognitive Abilities (WJ-III COG; R.W. Woodcock, K.S. McGrew, & N. Mather, 2001) were administered to 131 children in grades 3 through 5 who took part in a concurrent validity study included in the *Woodcock-Johnson Tests of Cognitive Abilities*, *Third Edition*, technical manual (K.S. McGrew & R.W. Woodcock, 2001). Confirmatory factor analyses using maximum likelihood estimation were conducted with the AMOS 5.0 (J.L. Arbuckle, 2001) statistical program to evaluate three models of increasing complexity, to compare how well each fit the data set, and to identify the one that best described the underlying constructs measured by the DAS. Results suggested that the synthesized Three-Stratum CHC Model provided the most parsimonious representation among the three models tested.

This study examined the underlying constructs measured by the Differential Ability Scales (DAS; Elliott, 1990a) as they relate to the Cattell-Horn-Carroll (CHC) Theory (McGrew, 1997) of cognitive abilities. The CHC Theory has become one of the most accepted contemporary theoretical models of cognitive abilities (Evans, Floyd, McGrew, & Leforgee, 2001; Keith, Kranzler, & Flanagan, 2001; McGrew, 1997). Because the latest version of the Woodcock-Johnson Tests of Cognitive Abilities (WJ-III COG; Woodcock, McGrew, & Mather, 2001) was designed to provide a comprehensive measure of the CHC Hierarchical Model, it is considered a psychometrically sound instrument both to analyze the underlying factor structure of the DAS and to identify the various CHC constructs or abilities assessed by the DAS battery (McGrew & Woodcock, 2001). Recent research advocates that a cross-battery approach be adopted to assess the major broad cognitive abilities adequately (Flanagan & McGrew, 1997; Flanagan, McGrew, & Ortiz, 2000; Flanagan & Ortiz, 2001). Consequently, a second objective of this study was to offer information regarding the unique contributions of the DAS in measuring narrow and broad abilities consistently with the CHC theoretical framework.

CHC Theory (McGrew, 1997) integrates three prominent theories of cognitive ability, Cattell's original Gf-Gc Theory (Cattell, 1971), Horn and Cattell's expanded Gf-Gc Theory (Horn & Cattell, 1966; Horn & Noll, 1997), and Carroll's Three-Stratum Model (Carroll, 1993). Cattell (1971) first proposed a theory distinguishing between two different components of *g*, or general ability, which he labeled as fluid (Gf) and crystallized (Gc) intelligence. He also discerned that fluid ability was influenced more by biological factors than was Gc and ultimately concluded that Gc was more related to educational and cultural factors. Horn expanded upon Cattell's original Gf-Gc Theory. This resulted in a total of nine broad ability factors, each of which consists of several narrower specific abilities (Cattell, 1963).

Correspondence to: David E. McIntosh, Department of Special Education, Teachers College 722, Ball State University, Muncie, IN 47306-1099. E-mail: jnlpsychscho@bsu.edu

Expanded Gf-Gc Theory (Horn & Cattell, 1966; Horn & Noll, 1997) became a popular conceptualization for understanding the structure of human intellectual abilities and processes (Bickley, Keith, & Wolfe, 1995; Cole & Randall, 2003; Dunham, McIntosh, & Gridley, 2002; Flanagan & McGrew, 1998; Flanagan et al., 2000; Keith, 1997; Woodcock, 1990; Ysseldyke, 1990). In 1993, John Carroll sought to unify the study of cognitive abilities by conducting a comprehensive analysis of psychometric research. His analysis included many of the studies investigating aspects of the Gf-Gc Model and resulted in the publication of *Human Cognitive Abilities: A Survey of Factor-Analytic Studies* (1993), deemed the most comprehensive view of existing literature on cognitive performance. As a result, Carroll described intellectual ability as a hierarchical structure where *g* exists as an overall, general ability, with eight second-order abilities located at the second stratum and numerous (approximately 69) specific abilities at the first-stratum level (Carroll, 1993; Brody, 2000). After analyzing the structure of Carroll's model and recognizing the considerable similarities between it and Gf-Gc Theory, McGrew proposed a cognitive model, called the CHC Theory of Cognitive Abilities, that integrated these prominent theories (McGrew, 1997).

The basic structure of McGrew's CHC Model defined human cognition as hierarchical in nature. His model retained Carroll's conceptualization of a three-tier organization of abilities with an overall general ability factor residing at the apex, 10 broad ability factors located at the second stratum, and numerous specific (Stratum I) abilities subsumed under each of the broad abilities (McGrew, 1997). However, it is important to note that McGrew (K. McGrew, personal communication, April 13, 2006) recognizes that the Cattell-Horn and Carroll models disagree on the presence of *g* and that he has stayed neutral on whether *g* actually exists; therefore, he typically tests models with *g* at the apex only because it is expected. McGrew also resolved some of the major discrepancies between the Horn-Cattell Gf-Gc Model and Carroll's model related to classification and inclusion/exclusion of certain broad and narrow abilities. On the basis of his own independent factor analyses of these models, McGrew identified 10 broad ability factors (versus Carroll's eight Stratum II factors), which were more reflective of Horn and Cattell's model (McGrew, 1997).

The Woodcock-Johnson Psycho-Educational Battery (WJTCA; Woodcock & Johnson, 1977) was originally developed in the absence of any dominant theoretical model of intelligence. However, this test provided the first comprehensive, co-normed battery of cognitive abilities and achievement skills, both areas that commonly are evaluated in psychoeducational assessments (Mather & Gregg, 2001). It also was one of the first batteries that included novel tasks not incorporated in preceding batteries (Kaufman, 2000). Inspired by Horn's (1985) expanded version of the Gf-Gc Theory, the 1989 revision of the Woodcock-Johnson Psycho-Educational Battery (WJ-R; Woodcock & Johnson, 1989) was designed to measure seven broad ability factors from Horn and Cattell's Theory. Two additional factors were measured on the achievement scale. The WJ-R was one of the most comprehensive batteries of its time and described as a "good translation of theory to practice" (Kaufman, 2000, p. 462). Furthermore, statistical analyses conducted on the WJ-R with other measures indicated that the Gf-Gc Theory could be applied to the interpretation of other ability tests (Schrank, Flanagan, Woodcock, & Mascolo, 2002). Still, the measure was criticized for using a limited number of individual tasks to define each broad ability factor.

Influenced by additional empirical research, the WJ-III COG was revised on the basis of the CHC Theory (McGrew, 1997). This revision incorporated Carroll's Hierarchical Stratum Theory with Horn and Cattell's Broad Ability Factor Theory to describe the structure of human cognitive abilities. Specifically, Schrank and Flanagan, 2003, stated that the "primary purpose of the WJ-III COG is to provide measurement of the broad CHC factor scores" (p. 6). Creating tests that measured the CHC abilities allowed for analysis of within-individual variability and provided additional ipsative interpretative information. Additionally, Woodcock proposed that his cognitive performance model be used to organize the broad ability factors and account for the influence that noncognitive factors have on an individual's performance (Mather & Gregg, 2001). However, given the considerable number of narrow abilities identified in the CHC Theory, no current battery, including WJ-III COG, has been acknowledged as adequately measuring all CHC abilities (Mather & Gregg, 2001).

Concurrently with the development of the WJ-R, Elliott developed an updated American version of his British Ability Scales (BAS, Elliott, 1983) to be used with preschool- and school-aged children. The resulting DAS was deliberately created in the absence of a specific theoretical model in order to make the battery more eclectic. Doing so allowed it to accommodate practitioners with a wide variety of theoretical perspectives (Elliott, 1990b). However, Elliott and others (e.g., Byrd & Buckhalt, 1991; Keith, 1990; Stone, 1992) have conducted studies that confirmed the structure of the DAS preschool- and school-age batteries as hierarchical in nature. The DAS has an overall general ability factor that comprises second-order ability clusters. Furthermore, Elliott provided evidence that ability becomes more differentiated with age.

Elliott's determination to make a battery that could be used as a diagnostic tool, allowing for interpretation of distinct abilities, was unique (Gordon & Elliott, 2001). Specifically, Elliott distinguished between core subtests, which factorially "load" onto higher-order ability clusters, and diagnostic subtests, which measure unique and relatively independent abilities that do not contribute to group factors (Elliott, 1990c). In addition, the administration of the DAS, during which the child is given a set of items appropriate his or her age level, was in direct contrast to more traditional methods used with other measures. Because the DAS appears compatible with the Horn-Cattell Gf-Gc Theory and is unique in its attempt to reflect specificity at the subtest level, it has been hypothesized that it should provide additional, unique descriptive information if interpreted from a multiple constructs theory of ability (e.g., the CHC Theory).

Given that no current ability measure has been found to assess all the broad and narrow cognitive abilities identified in the CHC Model (Carroll, 1993; McGrew, 1997; Woodcock, 1990), this study provides further clarification of the underlying factor structure of *two* measures of cognitive abilities. It explores the congruency and specificity of the factors measured by the WJ-III COG and DAS, which, in turn, may offer practitioners and researchers a greater breadth of coverage of the CHC framework. Researchers have suggested that practitioners "cross" batteries to ensure that sufficient range in the specific and broad abilities has been assessed, especially those abilities/domains shown to be related to academic achievement (Evans et al., 2001; Flanagan & McGrew, 1998; McGrew, Keith, Flanagan, & Vanderwood, 1997). Proponents of the cross-battery approach have recommended supplementing cognitive batteries that lack certain measures of the CHC constructs by administering additional, unique subtests from separate measures (Flanagan & McGrew, 1997). Ultimately, better classification of the abilities measured by various popular cognitive tests, such as the DAS, will help practitioners design a more comprehensive battery of tests. In turn, this will allow for the assessment of a broader range of abilities and the investigation of specific areas of cognitive functioning related to individual referral concerns (Flanagan & McGrew, 1997). It is anticipated that results from this study will clarify how well the DAS and WJ-III COG batteries, in combination, represent a "complete" assessment of the CHC Model. Moreover, by providing evidence that the DAS is a satisfactory measure of certain CHC abilities, these results will offer support for its continued use as an empirically validated assessment tool.

Methods

Participants

The participants were 131 children in grades 3 through 5 who took part in a concurrent validity study examining the DAS and WJ-III COG for inclusion in the *Woodcock-Johnson Tests* *of Cognitive Abilities–Third Edition* technical manual (McGrew & Woodcock, 2001). The sample included 69 males (52.7%) and 62 females (47.3%), of whom 125 were Caucasian, 5 African American, and 1 of Hispanic descent. Participants ranged in age from approximately 8 years, 3 months, to 12 years, 3 months (*M* = 120.67 months, or approximately 10 years, 1 month; *SD* = 10.50 months). Children were from public (91.6%) and private (8.4%) elementary schools located in south central Kentucky (Murray) and the metropolitan area of Albany, New York. The sample from Kentucky consisted of 42 males and 35 females, whereas the New York sample included 27 males and 27 females. There were 39 third graders, 49 fourth graders, and 43 fifth graders. None of the participants had ever received any type of special education or related services; however, four of the participants had repeated a grade. Participants' parental education levels were as follows: 50.4% of mothers and 47.3% of fathers (missing this information for one case) had completed at least 4 years of college, 29.0% of mothers and 20.6% of fathers had some college training, 19.1% of mothers and 25.2% of fathers had high school diplomas, and the remaining parents (1.5% of mothers, 6.1% of fathers) had completed less than a high school education. None of the children was reported to have any type of psychological or educational disorder according to the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV; American Psychiatric Association, 1994).

Instrumentation

Differential Ability Scales. The DAS consists of 17 different cognitive subtests divided into preschool- (from 2 years, 6 months, to 5 years, 11 months) and school-age- (6 years, 0 months, to 17 years, 11 months) level batteries, that allow for out-of-level testing procedures with exceptional children within the age range of 5 years to approximately 8 years (Elliott, 1990c). Although the two batteries are distinct, there is some overlap in terms of subtests included on both. The DAS also has three achievement subtests (Word Reading, Spelling, and Basic Number Skills) that can be administered along with the school-age cognitive battery.

Only the nine cognitive tests included on the school-age battery were of interest in the current study. Six core subtests make up the General Conceptual Ability (GCA) composite score. Two of the six core subtests contribute to each of the three cluster scores. Specifically, the Verbal Ability Cluster, which measures "complex verbal mental processing such as acquired verbal concepts, verbal knowledge, and reasoning" (Elliott, 1990c, p. 83), comprises Word Definitions and Similarities subtests. The Nonverbal Reasoning Cluster, which includes the Matrices and Sequential and Qualitative Reasoning subtests, reflects an individual's nonverbal, inductive reasoning abilities (Elliott, 1990c). Finally, the Spatial Ability Cluster, a measure of visual-spatial perceptual organization and analytical thinking abilities, consists of the Recall of Designs and Pattern Construction subtests (Elliott, 1990c). The remaining three subtests (Recall of Digits, Recall of Objects, and Speed of Information Processing) are diagnostic and do not contribute to any of the cluster scores. The standardized mean for the GCA and Cluster scores is 100, with a standard deviation of 15. Subtest scores have a standardized mean of 50 and standard deviation of 10.

Woodcock-Johnson Tests of Cognitive Abilities–Third Edition (WJ-III COG). The WJ-III COG is an individually administered measure of cognitive abilities designed to assess individuals from 2 to 90+ years of age (Woodcock, McGrew, & Mather, 2001). It consists of 20 distinct subtests. These are divided into the Standard Battery (seven standard and three supplemental subtests) and the Extended Battery (10 additional tests). Seven of the CHC ability factors are measured by the WJ-III COG: comprehension-knowledge (Gc), long-term retrieval (Glr), auditory processing (Ga), visual-spatial thinking (Gv), auditory processing (Ga), fluid reasoning (Gf), and processing speed (Gs). Not all of the CHC factors were included in the current study because not all subtests were administered during the standardization phase of the WJ-III COG. Time constraints precluded administering all tests. Therefore, tests that contributed to six of seven of the CHC ability factors were administered. None of the tests related to auditory processing (Ga) was given, because there is no empirical evidence suggesting that the DAS includes subtests of auditory processing (McGrew, 1997). WJ-III COG tests that theoretically corresponded to DAS subtests were chosen for the current study. Specifically, the Verbal Comprehension test, which contributes to the comprehension-knowledge (Gc) factor and visual-auditory learning (long-term retrieval, Glr), and the spatial relations and picture recognition tests (visual-spatial thinking, Gv) were administered. Concept formation and analysis-synthesis tests were included and composed the fluid reasoning (Gf) factor. Processing speed (Gs), which comprises Visual Matching and Decision Speed tests, also was given. Finally, Numbers Reversed, Memory for Words, and Auditory Working Memory tests (short-term memory, Gsm) were administered. The WJ-III COG provides a General Intellectual Ability (GIA) score (overall ability factor) that consists of a composite of the individual tests. All WJ-III COG scores (GIA score, Cluster cores, and individual test scores) are based on a standardized mean of 100 and standard deviation of 15.

Procedures

Participants were general education school-age children who had volunteered for the project. They were solicited through the students' teachers, who sent participant requests/consent forms home with each child. Parental permission (signed consent) forms, as well as child consent forms, were obtained for all participants. In addition, children received $10 for their participation in the study.

The complete DAS school-age battery and the selected tests from the WJ-III COG mentioned were individually administered to all participants by trained examiners. Examiners included both licensed school psychologists and advanced graduate students who had completed training in psychological assessment and who were supervised by doctoral-level school psychologists. In addition, all examiners completed a 2-day training workshop conducted by research directors from Riverside Publishing Company. All testing was completed after the school day ended. The DAS and WJ-III COG were administered in a counterbalanced order in one to two sessions. No more than 7 days elapsed between testing sessions. Total administration time for the two batteries was approximately 2.5 to 4.0 hours.

Results

Means, standard deviations, and ranges for the DAS General Conceptual Ability (GCA), Clusters (Verbal Ability, Nonverbal Reasoning, and Spatial Ability), and subtest scores are presented in Table 1. As a group, the participants of this study scored within the average range on the GCA ($M = 109.24$, $SD = 13.59$) and Ability Clusters (means ranging from 106.31 to 109.66). This was expected, given that the sample consisted of school-aged children participating in regular education curriculum.

The normality, or distribution of scores, was inspected to assess skewness and kurtosis values of the observed variables (tests) (Tabachnick & Fidell, 2001). Values closer to 0.0 indicate a normal distribution. The DAS subtests skewness statistics all fell between -1.0 and 1.0, indicating adequate dispersion of scores, with Word Definitions reflecting slight positive skewness (.97). With the exception of the Word Definitions subtest, all of the DAS variables had kurtosis statistics close to 0.0, ranging from $-.07$ to $-.38$. Word Definitions demonstrated a moderate kurtosis score (3.0). This indicates a significant peak in distribution around the median, which, however, was most likely a result of the specific sample used in this study, which consisted of children expected typically to perform within the average range.

Table 1
Means, Standard Deviations, and Ranges of the DAS GCA, Cluster, and Subtests Scores

DAS GCA and Clusters	*N*	*Mean*	*SD*	Minimum	Maximum
GCA	131	109.24	13.59	79	145
Verbal ability	131	106.31	15.62	75	151
Nonverbal reasoning ability	131	109.66	13.71	83	144
Spatial ability	131	107.23	13.57	75	150
DAS core subtests (*t* scores)					
Word definitions	131	53.20	10.24	28	100
Similarities	131	54.85	10.71	27	80
Recall of designs	131	53.94	9.85	26	78
Pattern construction	131	55.16	8.43	37	77
Matrices	131	54.41	8.93	36	80
Sequential and quantitative reasoning	131	57.56	9.17	37	80
DAS diagnostic subtests (*t* scores)					
Recall of objects (immediate)	131	52.68	8.53	34	74
Recall of digits	131	54.02	11.41	27	80
Speed of information processing	130	54.92	9.69	29	80

Note. DAS = Differential Ability Scales; GCA = General Conceptual Ability. The DAS GCA and Cluster scores have a standardized $M = 100$, $SD = 15$. The DAS subtests have a standardized $M = 50$, $SD = 10$.

The means, standard deviations, and ranges for the WJ-III COG GIA, CHC Cluster, and individual test scores are presented in Table 2. The WJ-III COG GIA score ($M = 107.03$, $SD = 14.01$) and CHC Cluster (Gv, Gf, Gs, and Gsm) mean scores all fell within the average range. Table 2 also shows the influence of missing data on specific individual tests that load into the configuration of the GIA, Gv, Gf, Gs, and Gsm scores. The WJ-III COG CHC Cluster mean scores ranged from 94.41 for the Gv cluster to 108.43 for the Gf cluster. Means and standard deviations across the WJ-III COG individual tests fell within the average range, with the lowest mean score at 93.16 (Picture Recognition) and the highest mean score at 109.21 (Verbal Comprehension).

The distributions of the WJ-III COG observed scores were examined for skewness and kurtosis. Although the WJ-III COG tests demonstrated adequate standard deviation values and ranges, some of the variables/tests demonstrated slight skewness (both positive and negative) in scores. However, with the exception of Rapid Picture Naming, none of the skewness statistics exceeded 1.0, thereby indicating acceptable dispersion of scores (Curran, West, & Finch, 1996; Tabachnick & Fidell, 2001). In addition, three WJ-III COG tests demonstrated nonnormal (exceeding 1.0) kurtosis scores. These tests were Rapid Picture Naming (3.48), Numbers Reversed (1.23), and Auditory Working Memory (3.13). These positive kurtosis statistics indicated the distributions of these tests were peaked at the median with few scores at the tail ends of the distributions. This finding might be expected given that this was a sample of general education students. However, the Rapid Picture Naming test demonstrated positive skewness and a nonnormal kurtosis score. Thus, the data collected for this WJ-III COG test do not approximate a normal distribution and have a restricted range, resulting in an underestimate of the variance of this test. However, given the fact that the other variables displayed acceptable dispersions and normal kurtosis, we believe the analyses are sound.

Table 2
Means, Standard Deviations, and Ranges of the WJ-III COG GIA, Clusters, and Tests

WJ-III COG GIA and Clusters	*N*	*Mean*	*SD*	Minimum	Maximum
General intellectual ability (GIA)	122	107.03	14.01	80	148
Visual processing (*Gv*)	123	94.41	16.95	53	132
Fluid reasoning (*Gf*)	131	108.43	13.95	74	151
Processing speed (*Gs*)	129	106.51	15.26	73	145
Short-term memory (*Gsm*)	131	105.66	14.19	68	161
WJ-III COG individual tests					
Verbal comprehension	131	109.56	14.05	82	160
Visual-auditory learning	131	104.60	13.87	58	137
Spatial relations	123	98.78	17.16	58	148
Picture recognition	123	93.16	13.42	52	118
Concept formation	131	106.55	13.92	76	144
Analysis-synthesis	131	109.21	14.08	73	152
Decision speed	129	105.63	15.64	73	143
Rapid picture naming	131	96.63	13.60	20	123
Numbers reversed	131	105.07	16.41	51	156
Auditory working memory	131	104.30	15.96	58	170
Memory for words	131	104.92	14.26	69	152

Note. WJ-III COG = Woodcock-Johnson Tests of Cognitive Abilities. WJ-III COG scores have a standardized $M = 100$, $SD = 15$.

Confirmatory Factor Analyses

Three models of increasing complexity were analyzed and compared to determine how well each fit the data set and to identify the one that best described the underlying constructs measured by the DAS in relation to CHC Theory. Confirmatory factor analyses using maximum likelihood estimation were conducted with the AMOS 5.0 (Arbuckle, 2001) statistical program. Several fit indices were used to examine how well each of the proposed models "fits" the current data, including the chi-square goodness-of-fit statistic, the comparative fit index (CFI), the parsimony-adjusted comparative fit index (PCFI), the Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA; Kline, 2005).

Results for each model are displayed as path diagrams (see Figures 1, 2, and 3.). Each diagram contains observed variables (subtests/tests), indicated by a rectangle, and unobserved/latent variables (factors), designated by circles or ellipses. In addition, small circles labeled with the letter *e* represent error variables. The letter and numbers assigned to each error variable were insignificant; variables and factors have been assigned an error term to reflect that all measures have some degree of variance unique to each variable. Single-headed arrows within the path diagrams represent a direct or "causal" relation between a variable and factor, with the value (path coefficient) assigned to each arrow quantifying how well that variable loads onto the corresponding factor. Gridley (2002) describes the nature of the obtained path coefficients (factor loadings) between latent and observed variables as causal, explaining that the coefficients are a way of quantifying the amount of variability of the observed variables caused or influenced by the underlying latent factors. The loadings are reported as standardized estimates (scaled from 0.0 to 1.0). Note that Model 2 includes path diagrams with double-headed arrows that show correlation coefficients between the two connected factors.

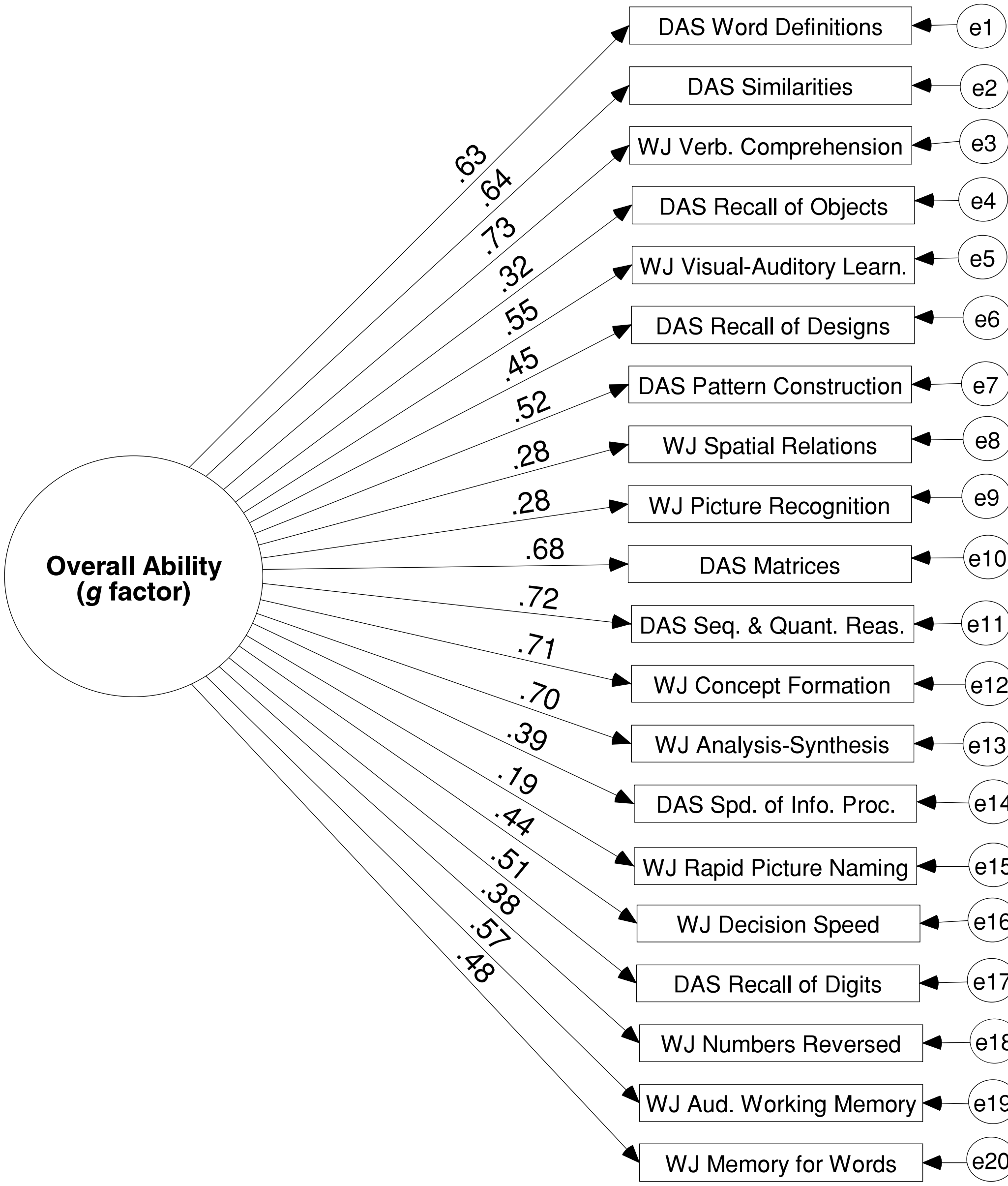

FIGURE 1. Model 1: One-Factor Model (Differential Ability Scales and Woodcock-Johnson Tests of Cognitive Abilities (WJ-III COG) test loading unto an overall *g* factor).

According to Comrey and Lee (1992), loadings of .71 and greater (indicating approximately 50% overlapping variance) are considered excellent; loadings between .63 and .70 are considered very good; .55 to .62 are considered good; .45 to .54 are considered fair; and .32 to .44 are considered poor. Tabachnick and Fidell (2001, p. 625) suggested only loadings of .32 and higher are interpretable. In addition to using these general guidelines, statistical significance tests for path coefficients can be obtained as a *t* value with corresponding probability statistic (*P* value). All path

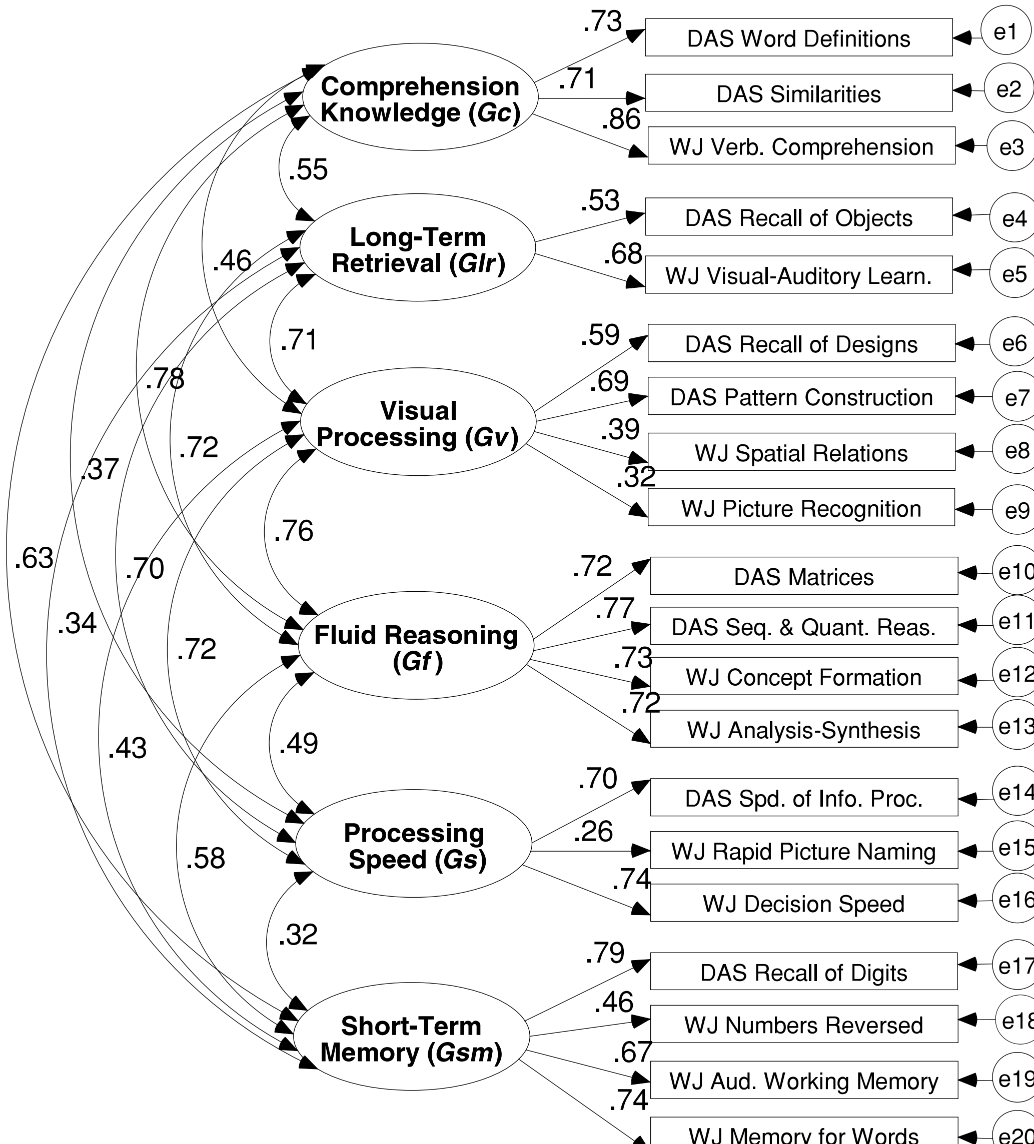

FIGURE 2. Model 2: Two-Stratum Cattell-Horn-Carroll (CHC) Model (with the Differential Ability Scales subtests and Woodcock-Johnson Tests of Cognitive Abilities [WJ-III COG] tests loaded unto six of the CHC cluster factors).

coefficients for the models examined in this study were statistically significant. Thus, the use of all observed variables in each respective model was supported.

Model 1 (One-Factor Model)

This One-Factor Model was developed as the "base" model to be compared with the more complex multifactor models, as described later. The path diagram with factor loadings of all subtests/tests for Model 1 is presented in Figure 1. The fit statistics for the One-Factor Model are

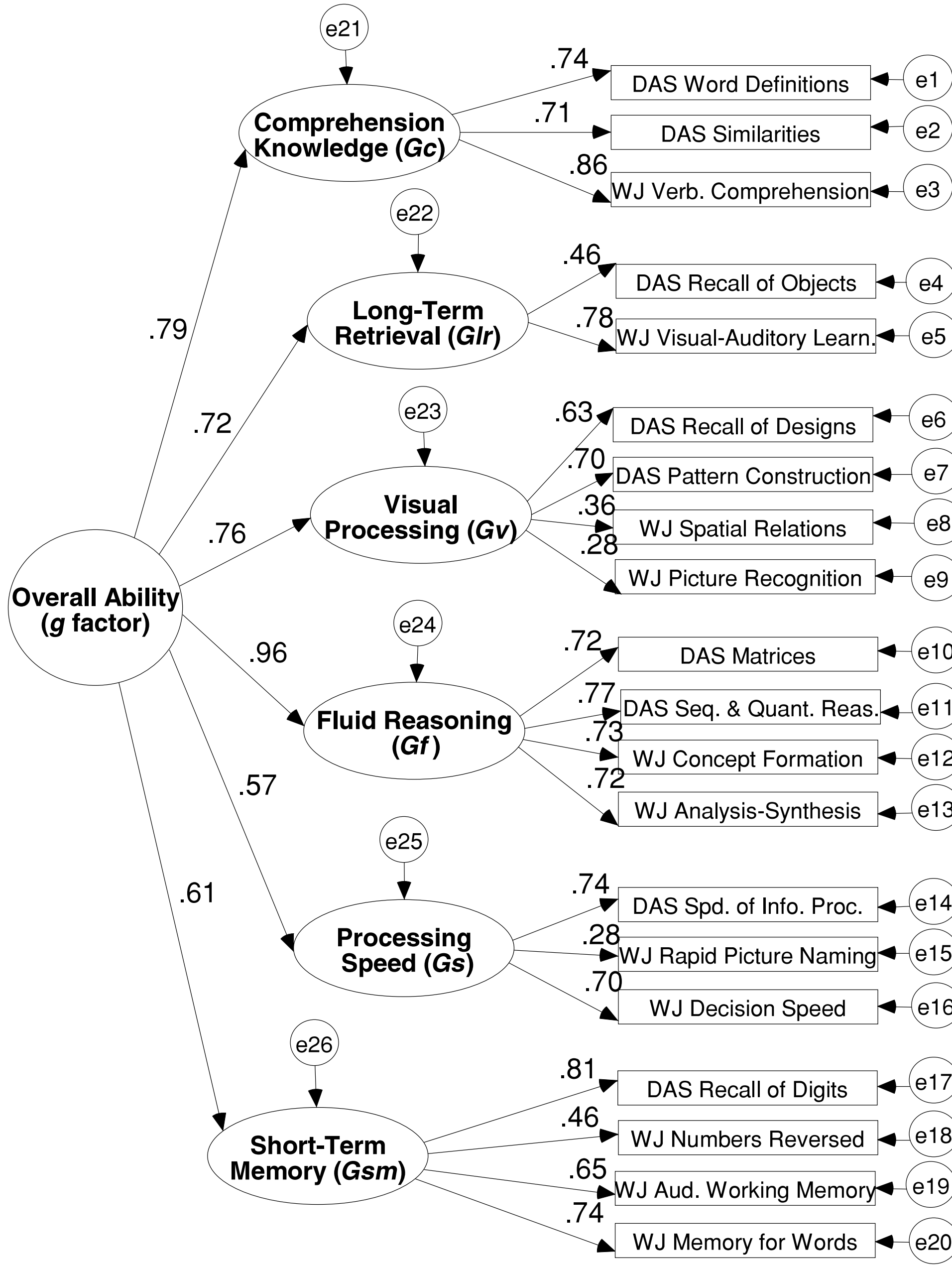

FIGURE 3. Three-Stratum Cattell-Horn-Carroll (CHC) Model (with six first-order CHC clusters and an overall *g* second-order factor).

displayed in Table 3. The chi-square ($\chi^2 = 372.52$, $df = 170$, $p = .000$) for Model 1 was significant, suggesting that a One-Factor Model (or implied covariance matrix from this sample) does not provide an adequate representation of the actual (true) model (or true covariance matrix). Chi-square is rarely used in isolation to determine the goodness of fit of a model because of its

Table 3
Fit Statistics for Models 1, 2, and 3

Models	$\chi^2(df)$	p	χ^2/df (ratio)	TLI	CFI	PCFI	RMSEA	AIC
Model 1 (One-Factor Model)	372.52 (170)	.000	2.19	.685	.745	.603	.096	492.517
Model 2 (Two-Stratum CHC Model)	195.93 (155)	.014	1.26	.930	.948	.700	.045	345.934
Model 3 (Three-Stratum CHC Model)	224.80 (164)	.001	1.37	.902	.923	.721	.053	356.804

Note. CHC = Cattell-Horn-Carroll; χ^2 = chi-square; (df) = degrees of freedom; p = probability; TLI = Tucker-Lewis Fit Index; CFI = Comparative Fit Index; PCFI = Parsimony-Adjusted Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; AIC = Aiken Information Criterion.

sensitivity to sample size and external variables such as sample distribution, a factor in this study. Examination of the remaining fit indices also suggests that the One-Factor Model was a poor fit.

Model 2 (Proposed Two-Stratum CHC Model)

The WJ-III COG was used in the development of this model as a validated representation of six of the CHC Stratum II broad factors. The WJ-III COG individual tests administered in this study were loaded onto their respective broad ability factors and used as markers for Stratum I and II CHC abilities. The DAS subtests were then loaded onto six CHC broad factors on the basis of a review of contemporary theoretical research (Carroll, 1993; Horn & Noll, 1997; Keith, 1997; McGrew, 1997; McGrew & Flanagan, 1998; McGrew & Woodcock, 2001) and analyzed in conjunction with the WJ-III COG tests. Correlations among the second-stratum factors were allowed to vary for this model.

The path diagram with factor loadings of all subtests/tests for this model is presented in Figure 2. The fit statistics for Model 2 are displayed in Table 3. As with Model 1, the chi-square statistic for the Two-Stratum CHC Model was significant (χ^2 = 195.93, df = 155, p = .014). Notably, this value was considerably lower than the chi-square statistic from Model 1 (χ^2 = 372.52, df = 170, p = .000). In addition, the χ^2 ratio of Model 2 was below the suggested criterion of 2.0 for an acceptable fit (Bollen, 1989). The remaining index values (*TLI* of .930, *CFI* of .948, *PCFI* of .700, and *RMSEA* of .045) were high and much improved when compared to those of Model 1. These fit scores indicated that the proposed Two-Stratum CHC theoretical model provided a good fit because Model 2 had a substantially lower Aiken Information Criterion (AIC) value and chi-square ratio than the One-Factor Model, thereby demonstrating a better representation of the data than Model 1. Finally, the test composed of the difference between the χ^2 values for Models 1 and 2 was significant, indicating that Model 2 had a better data-model fit than did Model 1 (see Table 4).

Examination of the factor loadings of each observed variable in Model 2 revealed that most subtests generally fit well with their respective factors. Specifically, all of the observed variables/measures loaded onto the comprehension-knowledge (Gc) and fluid reasoning (Gf) factors displayed excellent loadings on the basis of Comrey and Lee's (1992) guidelines. These results suggested that the DAS Word Definitions and Similarities subtests and WJ-III COG Verbal Comprehension test are good measures of these two CHC broad abilities. The two tests proposed as measures of long-term retrieval (Glr) demonstrated fair to very good loadings. Specifically, the DAS Recall of Objects subtest had a loading of .53. Of note, other researchers have described the DAS Recall of Objects subtest as a mixed indicator of memory. They have proposed that it could be classified as measuring both short-term memory and long-term retrieval (Keith, 1990;

Table 4
Chi-Square Comparison of Models 1, 2, and 3

Models	$\chi^2(df)$	$\Delta\chi^2$ (df)	*p*
Model 1	372.52 (170)		
Model 2	195.93 (155)	176.59 (15)[a]	<.000[a]
Model 3	224.80 (164)	147.72 (6)[a]	<.000[a]
		28.87 (9)[b]	<.000[b]

Note. χ^2 = chi square; (df) = degrees of freedom; $\Delta\chi^2$ = change in chi square; p = probability.
[a]Compared to Model 1 (One-Factor Model).
[b]Compared to Model 2 (Two-Stratum Cattell-Horn-Carroll Model).

McGrew, 1997; McGrew & Flanagan, 1998), and that characterization may account for its moderate path coefficient value. The processing speed (Gs) and short-term memory (Gsm) subtest/test path coefficients ranged in strength. Both second-stratum CHC factors had at least two very strong indicators/measures (loadings ranging from .70 to .79). However, these factors also had WJ-III COG tests with weaker loadings on their designated CHC factor. For example, WJ-III COG Numbers Reversed only displayed a "fair" loading on the Gsm factor.

Similarly, the Rapid Picture Naming test, one of the WJ-III COG supplementary tests, had a path coefficient/loading of .26 on the Gs factor. This suggests that this test is not a strong measure of processing speed and may better represent a different CHC ability. For example, it may also measure memory, given that this timed task requires the individual to recognize and identify objects. However, it is more likely that this lower path coefficient stems from the test's restricted range/limited variance as reflected in its nonnormal kurtosis and skewness statistics. As mentioned, Rapid Picture Naming only exhibited one significant, albeit modest, intercorrelation with the other individual WJ-III COG tests and DAS subtests, with DAS Speed of Information Processing subtest ($r = .274$). This result supports the placement of Rapid Picture Naming on the Gs CHC factor, although it was not a very strong measure of processing speed with this specific sample. Another consideration is that the low loading demonstrated by Rapid Picture Naming is due to the fact that it is the sole indicator in the model for an ability that is not represented in the current model. For example, Rapid Picture Naming has been found to load with Retrieval Fluency to form the factor Naming Facility (K. McGrew, personal communication, April 13, 2006).

The visual processing (Gv) factor had few strong indicators, revealing subtest/test loadings that ranged from .32 (poor) to .69 (very good). In fact, the DAS Recall of Designs and Pattern Construction subtests were considered better measures of Gv than the WJ-III COG tests (Spatial Relations and Picture Recognition) that were used as markers of the CHC Gv ability for this study. Notably, results of an exploratory factor analysis of the WJ-III COG, Wechsler Adult Intelligence Scale, Third Edition (WAIS-III), and Kaufman Adolescent and Adult Intelligence Test (KAIT) reported in the Woodcock-Johnson III technical manual (McGrew & Woodcock, 2001) indicated moderate instability of the WJ-III COG Gv cluster. Specifically, Spatial Relations separated into a secondary Gv factor with a KAIT subtest and Picture Recognition loaded onto Glr. Moreover, stronger intercorrelations were found between the DAS subtests purported to measure Gv and the WJ-III COG Gf tests (Concept Formation and Analysis-Synthesis) than with the WJ-III COG Gv tests in this sample. Similarly, the WJ-III COG Gv tests had stronger correlations with other DAS and WJ-III COG tests than with the DAS Recall of Designs and Pattern Construction subtests (measures of Gv). McGrew and

Woodcock (2001) reported that the WJ-III COG Gv cluster had several moderate factor loadings. They proposed that having several tests with high factor loadings might reflect limited factor breadth, suggesting the possibility that these tests measure the same narrow ability. This characteristic would obviously compromise the interpretation of the construct as a broad ability factor. In other words, the test authors explained that having a greater number of tests with moderate to strong loadings implied that each test is measuring a unique aspect of the Gv primary ability. So, in the context of this study, the fact that the loadings associated with Gv are somewhat lower than might be desired when considering Comrey and Lee's (1992) guidelines might actually suggest that in some sense Gv is a more validly measured construct because it is broadly defined.

Model 3 (Proposed Three-Stratum CHC Model)

This proposed model maintained many of the same properties as Model 2; however, a second-order general ability factor (Stratum III) was added to reflect the complete hierarchical CHC Theory of cognitive abilities. This model has three different levels, or tiers, with the individual subtests/tests (observed variables) of the DAS and WJ-III serving as indicators of the CHC narrow (Stratum I) abilities; the first-order latent CHC factors (Gc, Glr, Gv, Gf, Gs, and Gsm) representing the Stratum II broad abilities; and the overall *g* factor at the apex of this hierarchical model.

The path diagram with first-order and second-order factor loadings of the subtests/tests of the Three-Stratum CHC Model is presented in Figure 3. The fit statistics for Model 3 are displayed in Table 3. As with the other models, the chi-square statistic for the Three-Stratum CHC Model was significant ($\chi^2 = 224.80$, $df = 164$, $p = .001$). With the exception of the chi-square, the χ^2 ratio (1.37) and remaining index values ($TLI = .902$, $CFI = .923$, $RMSEA = .053$) for Model 3 suggested that this model provided a fairly good fit.

The second-order *g* loadings for the DAS subtests and WJ-III COG tests are indicated in Figure 3. The CHC factor loadings on *g* were substantial, ranging from .57 for processing speed (Gs) to .96 for fluid reasoning (Gf). These path coefficients indicated that an overall general ability factor does, in fact, account for a substantial amount of variance of CHC broad abilities. This finding verifies McGrew's retention of an underlying general ability factor in his CHC theoretical model.

The majority of the direct effects/path coefficients of the first-order CHC factors on the individual test variables were large; only 7 of the 20 configured standardized path coefficients (factor loadings) fell below .70. Essentially, the factor loadings of this hierarchical model were equivalent to the correlated Two-Stratum CHC Model. Nevertheless, moderate increases in loadings were evidenced on several subtests/tests in the Three-Stratum CHC Model. For example, the WJ-III Visual-Auditory Learning loading on long-term retrieval (Glr) improved from .68 in Model 2 to .78 in Model 3; the DAS Speed of Information Processing loading on processing speed (Gs) improved from .70 in Model 2 to .74 in Model 3; and the two DAS visual processing (Gv) variable loadings (Recall of Designs and Pattern Construction) increased. As with the Two-Stratum CHC Model, the WJ-III COG Spatial Relations, Picture Recognition, and Rapid Picture Naming tests continued to have low loadings on their respective factors (Gv and Gs). The DAS subtest loadings were examined specifically, given that the focus of this study was to determine whether the DAS measured the CHC constructs. All of the DAS subtests demonstrated "excellent" loadings on their respective CHC factors with the exception of Recall of Objects on Glr (.46 = fair) and Recall of Designs on Gv (.63 = very good). Again, these results supported the interpretation of the DAS from the CHC theoretical perspective.

Indirect effects of the second-order *g* factor on the DAS subtests and WJ-III COG tests also were calculated by multiplying the first-order and second-order factor loadings of each of the individual tests (Kline, 2005). These values appear in Table 5. According to Kaufman's (1994)

Table 5
Second-Order Factor Loadings of the DAS and WJ-III COG Tests on the Three-Stratum

Subtest/Test	Second-order loadings *g* (Stratum III)
DAS	
Word definitions	.59
Similarities	.56
Recall of designs	.48
Pattern construction	.53
Matrices	.69
Sequential and quantitative reasoning	.74
Recall of objects	.33
Recall of digits	.49
Speed of information processing	.42
WJ–III COG	
Verbal comprehension	.68
Visual-auditory learning	.56
Spatial relations	.27
Picture recognition	.21
Concept formation	.70
Analysis-synthesis	.69
Rapid picture naming	.16
Decision speed	.40
Numbers reversed	.28
Auditory working memory	.40
Memory for words	.45

classification of *g* loadings, values of .70 or greater indicate a good measure of *g*; values of .50 to .69 indicate a fair measure of *g*; loadings lower than .50 are deemed poor measures of *g*. It should be noted that these guidelines are based upon the use of principal components analysis (PCA) to obtain the loadings. Research has shown that loadings obtained from PCA tend to be somewhat larger than those obtained from principal axis or other factor methods (Thompson, 2004). Therefore, while these guidelines must be used with some care when employed with the results from the CFA presented here, they will be somewhat conservative. In other words, the loadings obtained from CFA are somewhat lower than would be those obtained from a PCA. Therefore, the Kaufman guidelines might be seen as somewhat stringent, and the true strength of the *g* loadings might be somewhat greater than is reported here. However, there is not a readily available methodology for employing the PCA extraction method in the confirmatory context. For this reason, we will rely on the results from the CFA and interpret them by using Kaufman's guidelines, keeping in mind the somewhat conservative nature of this approach.

Only two tests, the DAS Sequential and Quantitative Reasoning subtest and the WJ-III COG Concept Formation test, demonstrated a good *g* loading. This result was not surprising because both tests were considered to be strong measures of Gf, which is the CHC broad ability that has been found to correlate highest with *g* (Keith, Quirk, Schartzer, & Elliott, 1999). In addition, the second-order *g* loadings for the other Gf subtests/tests, DAS Matrices (.69) and WJ-III COG Analysis-Synthesis (.69), approached the .70 criterion. Word Definitions, Similarities, Pattern Construction on the DAS, and Verbal Comprehension and Visual-Auditory Learning tests of the

WJ-III COG also demonstrated fair second-order g loadings. The remaining g loadings were poor. Of interest, the g loading for the DAS Recall of Digits (.49) was higher than expected, given that it is one of the diagnostic subtests that were considered to be poor measures of g (Elliott, 1990c).

To determine which of the three proposed models offered the best fit to the data, the χ^2 ratio and AIC statistics were examined, with lower values of each suggesting superior fit. Moreover, the chi-square differences, or changes in chi-square, across the three nested models were tested for significance (see Table 4). As predicted, both Model 2 and Model 3 demonstrated significant improvements in fit compared to the One-Factor Model (Model 1 χ^2 ratio = 2.19, AIC = 492.517; Model 2 χ^2 ratio = 1.26, AIC = 345.934; Model 3 χ^2 ratio = 1.37, AIC = 356.804). This result indicated that a one-factor theoretical model was inferior to a multifactorial theory. Furthermore, the difference between the chi-square values of the models displayed in Table 4 also confirmed that the Two-Stratum CHC Model fits significantly better than the One-Factor Model and the Three-Stratum CHC Model.

When comparing Models 2 and 3, the difference between the chi-square ratios, AIC, TLI, and CFI fit statistic values of the Two-Stratum CHC Model and Three-Stratum CHC Model was minimal; however, on the basis of the χ^2 difference test discussed earlier, the Three-Stratum hierarchical model reflected a poorer fit to the data. This decrease in fit for the Three-Stratum CHC Model is not an uncommon occurrence when higher-order factors are added to a model to account for correlations among first-order factors (Gorsuch, 1983). Specifically, higher-order factors, such as the g factor in Model 3, sometimes reduce the accuracy of a model because the addition of hierarchical factors actually increases the breadth of generalization by accounting for correlations evidenced between factors. This contrasts with a model of first-order factors (Model 2) that focuses more on describing and interpreting the common variance evidenced between narrow/specific observed variables (Gorsuch, 1983). Although the Two-Stratum CHC Model, which is similar to the model (without hierarchical g) used in the development of the WJ-III COG, provided the most probable fit to the data as evidenced by better-fitting index statistics, the differences were not necessarily large enough to distinguish Model 2 as more tenable than Model 3. Moreover, the addition of an overall g factor score has significant practical use in the field of cognitive assessment, given that many clinicians use an overall global score for making diagnostic decisions.

Discussion

Model 1 was designed to reflect a single-factor solution, with all DAS subtests and WJ-III COG tests specified as measures of g. Model 1 was proposed as a baseline model for comparing the more complex models and was expected to be a relatively poor representation of the data given that the DAS and WJ-III COG were designed to measure multiple abilities. As predicted, this study uncovered an inadequate fit to the data and, therefore, did not find ample support for Model 1.

Examination of the One-Factor Model revealed that the highest loadings were measures of comprehension-knowledge (WJ-III COG Verbal Comprehension) and fluid reasoning (DAS Sequential and Quantitative Reasoning, WJ-III Concept Formation, and WJ-III Analysis-Synthesis). These results were similar to those found in McGhee's (1993) study of the DAS, the Woodcock-Johnson Psycho-educational Battery-Revised (Woodcock & Johnson, 1989), and the Detroit Test of Learning Aptitude, Third Edition (Hammill & Bryant, 1991), wherein tests measuring verbal/lexical comprehension ability (or Gc) and fluid reasoning (or Gf) demonstrated some of the highest path coefficients on g (e.g., DAS Word Definitions loading of .847, DAS Similarities loading of .821, WJ-R Concept Formation loading of .723, WJ-R Analysis-Synthesis loading of .734). The lowest path coefficients were from measures of memory (DAS Recall of Objects) and visual processing (WJ-III COG Spatial Relations and Picture Recognition). These results were not surprising given that research has identified g as having only moderate effects on these broad ability domains.

Moreover, the DAS diagnostic subtests had poor to fair loadings on *g*. This result was consistent with Elliott's decision to designate these subtests as distinct measures of specific abilities and exclude them from the core group of tests used to configure the DAS cluster and overall composite scores (Elliott, 1990c). WJ-III COG Rapid Picture Naming demonstrated the weakest loading, which may be an artifact of its nonnormal distribution or more likely may indicate that it was the sole indicator of an ability (Naming Facility) that was not reflected in the current model.

Next, the Two-Stratum CHC Model (Model 2) was examined to determine how well a contemporary Gf-Gc six-factor solution fit the data. This model included six of the CHC broad ability factors, with the exception of the auditory processing factor (Ga), with each factor having two to four of the DAS and WJ-III COG tests specified as primary measures.

Among the DAS subtests, Word Definitions and Similarities subtests were theorized as primary measures of comprehension-knowledge (Gc); Recall of Objects, as an indicator of long-term retrieval (Glr); Recall of Designs and Pattern Construction subtests, as measures of visual processing (Gv); Matrices and Sequential and Quantitative Reasoning, as measures of fluid reasoning (Gf); Speed of Information Processing, as a test of processing speed (Gs); and Recall of Digits, as a measure of short-term memory (Gsm). These specifications were based on past research of empirical and logical analyses of the constructs measured by the DAS cognitive battery (Flanagan & Ortiz, 2001; Keith, 1990; McGrew, 1997; Stone, 1992). Moreover, Elliott (1997) conceded that the DAS factor structure could be interpreted from the Gf-Gc Theoretical Model. The WJ-III COG tests were specified as markers of their respective CHC factors as indicated by the test authors.

Ultimately, the six-factor solution provided an excellent representation of the constructs measured by the DAS. Each of the DAS subtests was designated as a good to excellent measure of its respective CHC ability factor. These results supported the interpretation of the DAS subtests as primary measures of the CHC Stratum II abilities (McGrew & Flanagan, 1998; Flanagan & Ortiz, 2001). Moreover, the DAS subtests' loadings on the Gc, Gf, and Gv factors were comparable to those reported by McGhee for his Seven-Factor Gf-Gc Model. However, McGhee's specification of DAS subtests was not congruent with some of the specifications made for Model 2, given that his research predates McGrew's classification of tests from the synthesized CHC theoretical perspective. Notably, the WJ-III COG Gv tests demonstrated weak loadings on the CHC Gv factor when specified as indicators of Gv with the DAS Recall of Designs and Pattern Construction subtests. This result indicated that these four tests likely measure unique, specific abilities.

Finally, a synthesized Three-Stratum CHC Model (Model 3) was developed to represent the entire hierarchical CHC Theory. This factorial model was identical to Model 2, with the addition of a higher-order general ability (*g*) factor to account for the correlations evidenced between the Stratum II broad abilities. In addition, this Three-Stratum Model provided information regarding the indirect, unique contributions of *g* on the individual DAS and WJ-III COG tests.

First-order direct effects of the CHC Stratum II abilities on the DAS and WJ-III COG tests in Model 3 were comparable to those found in Model 2, although there were slight, inconsequential variations in loadings for some of the individual tests (WJ-III COG Visual-Auditory Learning, WJ-III Spatial Relations, DAS Recall of Objects, and DAS Speed of Information Processing). The influence of *g* on the CHC broad abilities was substantial for all six factors. However, with the exception of the Gf, several of the path coefficients were only moderate in strength, supporting McGrew's and other researchers' (Carroll, 1993; Keith, 1990; Keith et al., 1999) conclusions that this smaller group of CHC broad factors measures different dimensions of cognitive ability. Fluid reasoning had a high loading on *g*. This finding corroborates previous research suggesting that measures of Gf are strongly related to overall *g* (Keith, 1990; Horn & Noll, 1997). In general, the indirect influence (second-order *g* loadings) of *g* on the individual DAS and WJ-III COG tasks generally was moderate. These data confirmed that the DAS and WJ-III COG individual tests do, in fact, provide additional interpretive information regarding an individual's cognitive functioning across different types of abilities other than *g*.

Comparison of the three models revealed that both multifactor models (Models 2 and 3) provided a much better solution in describing the constructs measured by the DAS and WJ-III COG than the One-Factor Model, as evidenced by the substantial improvements in the fit statistics. The Two-Stratum CHC Model and Three-Stratum CHC Model both offered equally tenable explanations of the DAS factor structure, although each theoretical model has its advantages. The Two-Stratum Model had the best fit across most of the goodness-of-fit statistics examined, thereby offering valuable empirical validity of the utility of the DAS when conducting CHC cross-battery assessments. Theoretically, it appears both the Two- and Three-Stratum Models are equally plausible from a CFA perspective and the Three-Stratum Model is not superior to the Two-Stratum Model; however, clinicians may find the Three-Stratum Model more practical. For example, psychologists and educators frequently require a global, or overall ability, score for making diagnostic and classification decisions, as is provided in the Three-Stratum CHC Model.

On the basis of this study, the DAS Word Definitions and Similarities subtests clearly resemble strong measures of comprehension-knowledge, thereby confirming Elliott's (1990c) description of these tests as indicators of verbal ability. Additionally, the DAS Matrices and Sequential and Quantitative Reasoning subtests should be interpreted as measures of fluid reasoning, and the DAS Speed of Information Processing and Recall of Digits diagnostic subtests may both be considered excellent measures of their respective CHC ability domains (processing speed and short-term memory). Caution should be exercised, however, when interpreting an individual's performance on the Information Processing and Recall of Digits diagnostic subtests separately because the DAS does not contain the suggested number of indicators (two or more qualitatively different narrow ability measures) to be considered an adequate representation of the processing speed and short-term memory CHC domains.

Recall of Objects on the DAS proved to be a fairly good predictor of long-term retrieval with a standardized coefficient of .54 found in the revised model. This weaker relation may reflect the fact that this task most likely is a mixed measure of both long-term retrieval and short-term memory. The individual is asked to recall previously presented pictures several times throughout the testing procedure with increasing time elapsing between each recall. McGrew (1997) classified the DAS Recall of Objects subtest as a possible measure of both Glr and Gsm on the basis of the logical analysis of the nature of this task. These results imply that psychologists should exercise caution when interpreting this subtest. Also, the Recall of Objects subtest may be best viewed as a mixed, or more generalized, measure of memory rather than of long-term retrieval.

As mentioned earlier, the DAS Gv tests (Recall of Designs and Pattern Construction) appeared to be very good measures of visual processing; however, these subtests most likely measure different narrow abilities than the WJ-III COG Gv indicators, given that the WJ-III COG tests exhibited fairly poor path coefficients when combined with these DAS subtests. Another implication of these findings is that the WJ-III COG Picture Recognition and Spatial Relations subtests do not appear to be robust measures of Gv when crossed with the DAS battery. As mentioned, this discovery is similar to findings of joint CFA research that has been conducted with the WJ-R but was unexpected given that the newer, revised WJ-III COG battery was used. Again, those conducting a cross-battery assessment with the WJ-III COG and the DAS, and possibly any other cognitive measure, should interpret results of the WJ-III COG Gv narrow ability tests with caution.

Moderate to strong intercorrelations were found between the Gc tests and broad Gc factor with other CHC tests and factors (Gsm and Gf). These results indicate children tend to rely on verbal or Gc abilities, such as verbal mediation or language skills, to understand and solve tasks that appear to measure different abilities.

From a practical standpoint, there is some advantage in using the complete hierarchical CHC model when interpreting results from these two batteries. It is still common practice for clinicians to use an overall global "IQ" score when making diagnostic and educational decisions (e.g., qualification for special education services or qualification for social service assistance). Therefore, the inclusion of an overall general ability score makes sense. Thus, choosing which model to use would likely depend more on the type of assessment being conducted (cross-battery assessment versus tradition/brief assessment of global ability). Nevertheless, incorporating the DAS in a CHC cross-battery methodology would likely provide a valid as well as more comprehensive interpretation of the DAS than the three-factor solution originally proposed by the DAS author.

Future CFA studies should attempt to replicate these results with larger samples. Although the results of this study evidenced highly factor-saturated variables (as indicated by the size of the factor loadings), thereby indicating that the current results would likely be replicated in corresponding studies, it would be important to replicate these results with other samples. Future studies may wish to consider studying the DAS in relation to alternative models based on other theoretical conceptualizations of cognitive ability (e.g., the Planning, Attention, Simultaneous and Successive Model). Other models also may provide plausible explanations for interpretation of the DAS cognitive battery.

References

American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Arbuckle, J.L. (2001). AMOS users' guide 5.0. Chicago: SmallWaters.

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(3), 309–328.

Bollen, K.A. (1989). Structural equations with latent variables. Hoboken, NJ: Wiley-InterScience.

Brody, N. (2000). History of theories and measurements of intelligence. In R. Sternberg (Ed.), Handbook of intelligence (pp. 16–33). New York: Cambridge University Press.

Byrd, P.D., & Buckhalt, J.A. (1991). A multitrait-multimethod construct validity study of the Differential Ability Scales. Journal of Psychoeducational Assessment, 9, 121–129.

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

Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 54, 1–22.

Cattell, R.B. (1971). Abilities: Their structure, growth, and action. Boston: Houghton Mifflin.

Cole, J.C., & Randall, M.K. (2003). Comparing the cognitive ability models of Spearman, Horn and Cattell, and Carroll. Journal of Psychoeducational Assessment, 21, 160–179.

Comrey, A.L., & Lee, H.B. (1992). A first course in factor analysis (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Curran, P.J., West, S.G., & Finch, J.F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychological Methods, 1(1), 16–29.

Dunham, M., McIntosh, D.E., & Gridley, B.E. (2002). An independent factor analysis of the Differential Ability Scales. Journal of Psychoeducational Assessment, 20, 152–163.

Elliott, C.D. (1983). The British Ability Scales. Windsor, England: NFER-Nelson.

Elliott, C.D. (1990a). Differential Ability Scales. San Antonio, TX: Psychological Corporation.

Elliott, C.D. (1990b). The nature and structure of children's abilities: Evidence from the Differential Ability Scales. Journal of Psychoeducational Assessment, 8, 376–390.

Elliott, C.D. (1990c). Differential Ability Scales: Introductory and technical handbook. San Antonio, TX: Psychological Corporation.

Elliott, C.D. (1997). The Differential Ability Scales. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (pp. 183–208). New York: Guilford Press.

Evans, J.J., Floyd, R.G., McGrew, K.S., & Leforgee, M.H. (2001). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. School Psychology Review, 31(2), 246–262.

Flanagan, D.P., & McGrew, K.S. (1997). A cross-battery approach to assessing and interpreting cognitive abilities: Narrowing the gap between practice and cognitive science. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (pp. 314–325). New York: Guilford Press.

Flanagan, D.P., & McGrew, K.S. (1998). Interpreting intelligence tests from contemporary Gf-Gc theory: Joint confirmatory factor analysis of the WJ-R and KAIT in a non-white sample. Journal of School Psychology, 36(2), 151–182.

Flanagan, D.P., McGrew, K.S., & Ortiz, S.O. (2000). The Wechsler intelligence scales and Gf-Gc theory: A contemporary approach to interpretation. Needham Heights, MA: Allyn & Bacon.

Flanagan, D.P., & Ortiz, S.O. (2001). Essentials of cross-battery assessment. New York: John Wiley & Sons.

Gordon, B., & Elliott, C.D. (2001). Assessment with the Differential Ability Scales. In J.J.W. Andrews, D.H. Saklofske, & H.L. Janzen (Eds.), Handbook of psychoeducational assessment: Ability, achievement, and behavior in children (pp. 65–101). San Diego, CA: Academic Press.

Gorsuch, R.L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Gridley, B.E. (2002). In search of an elegant solution: Reanalysis of Plucker, Callahan, and Tomchin, with respects to Pyryt and Plucker. Gifted Child Quarterly, 46(3), 224–234.

Hammill, D.D., & Bryant, B.R. (1991). Detroit Test of Learning Aptitude–Third Edition: Examiner's manual. Circle Pines, MN: AGS.

Horn, J.L. (1985). Remodeling old models of intelligence. In B.B. Wolman (Ed.), Handbook of intelligence: Theories, measurements, and applications (pp. 267–300). New York: John Wiley & Sons.

Horn, J.L., & Cattell, R.B. (1966). Refinement of the theory of fluid and crystallized general intelligences. Journal of Educational Psychology, 57, 253–270.

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

Kaufman, A.S. (1994). Intelligent testing with the WISC-III. New York: Wiley.

Kaufman, A.S. (2000). Tests of intelligence. In R. Sternberg (Ed.), Handbook of intelligence (pp. 445–476). New York: Cambridge University Press.

Keith, T.Z. (1990). Confirmatory and hierarchical confirmatory analysis of the Differential Ability Scales. Journal of Psychoeducational Assessment, 8, 391–405.

Keith, T.Z. (1997). Using confirmatory factor analysis to aid in understanding the constructs measured in 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 Press.

Keith, T.Z., Kranzler, J.H., & Flanagan, D.P. (2001). What does the Cognitive Assessment System (CAS) measure? Joint confirmatory factor analysis of the CAS and the Woodcock-Johnson Tests of Cognitive Ability (3rd ed.). School Psychology Review, 30(1), 89–119.

Keith, T.Z., Quirk, K.J., Schartzer, C., & Elliott, C.D. (1999). Construct bias in the Differential Ability Scales? Confirmatory and hierarchical factor structure across three ethnic groups. Journal of Psychoeducational Assessment, 17, 249–268.

Kline, R.B. (2005). Principles and practice of structural equation modeling. New York: Guilford Press.

Mather, N., & Gregg, N. (2001). Assessment with the Woodcock-Johnson III. In J.J.W. Andrews, D.H. Saklofske, & H.L. Janzen (Eds.), Handbook of psychoeducational assessment: Ability, achievement, and behavior in children (pp. 133–165). San Diego, CA: Academic Press.

McGhee, R. (1993). Fluid and crystallized intelligence: Confirmatory factor analyses of the Differential Abilities Scale, Detroit Tests of Learning Aptitude–3, and Woodcock-Johnson Psycho-educational Battery–Revised. In R.S. McCallum & B.A. Bracken (Eds.), Woodcock-Johnson Psycho-educational Battery—Revised: A monograph of the Journal of Psychoeducational Assessment. Knoxville, TN: Psychoeducational Corporation.

McGrew, K.S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (pp. 151–182). New York: Guilford Press.

McGrew, K.S., & Flanagan, D.P. (1998). The intelligence tests desk reference (ITDR): Gf-Gc cross-battery assessment. Needham Heights, MA: Allyn & Bacon.

McGrew, K.S., Keith, T.Z., Flanagan, D.P., & Vanderwood, M. (1997). Beyond g: The impact of the Gf-Gc specific cognitive abilities research on the future use and interpretation of intelligence tests in the schools. School Psychology Review, 26(2), 189–210.

McGrew, K.S., & Woodcock, R.W. (2001). Technical manual: Woodcock-Johnson III. Itasca, IL: Riverside.

Schrank, F.A., & Flanagan, D.P. (2003). WJ-III clinical use and interpretation: Scientist-practitioner perspectives. San Diego, CA: Academic Press.

Schrank, F.A., Flanagan, D.P., Woodcock, R.W., & Mascolo, J.T. (2002). Essentials of WJ-III cognitive abilities assessment. Hoboken, NJ: John Wiley & Sons.

Stone, B.J. (1992). Joint confirmatory analyses of the DAS and WISC-R. Journal of School Psychology, 30, 185–195.

Tabachnick, B.G., & Fidell, L.S. (2001). Using multivariate statistics (4th ed.). Needham Heights, MA: Allyn & Bacon.

Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association.

Woodcock, R.W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. Journal of Psychoeducational Assessment, 8, 231–258.

Woodcock, R.W., & Johnson, M.B. (1977). Woodcock-Johnson PsychoEducational Battery. Allen, TX: DLM/Teaching Resources.

Woodcock, R.W., & Johnson, M.B. (1989). Woodcock-Johnson Tests of Cognitive Ability–Revised. Chicago: Riverside.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). Woodcock-Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside.

Ysseldyke, J.E. (1990). Goodness of fit of the Woodcock-Johnson Psychoeducational Battery–Revised to the Horn-Cattell Gf-Gc theory. Journal of Psychoeducational Assessment, 8, 268–275.