Journal of Psychoeducational Assessment
Volume 24 Number 4
December 2006 303-317

10.1177/0734282906287823
http://jpa.sagepub.com
hosted at
http://online.sagepub.com

Relations Between the Woodcock-Johnson III Clinical Clusters and Measures of Executive Functions From the Delis-Kaplan Executive Function System

Randy G. Floyd
Allison C. McCormack
Elizabeth L. Ingram
Amy E. Davis
Renee Bergeron
University of Memphis

Gloria Hamilton
Middle Tennessee State University

This study examined the convergent relations between scores from four clinical clusters from the Woodcock-Johnson III Tests of Cognitive Abilities (WJ III) and measures of executive functions using a sample of school-aged children and a sample of adults. The WJ III clinical clusters included the Working Memory, Cognitive Fluency, Broad Attention, and Executive Processes clusters, and the measures of executive functions were from the Delis-Kaplan Executive Function System (D-KEFS). Across both samples, all clinical clusters demonstrated evidence of statistically significant and moderate positive relations with at least some measures of executive functions. The Executive Processes cluster demonstrated relations with measures of executive functions that tended to be the strongest and most consistent of the WJ III clinical clusters. When these relations between the clinical clusters and the measures of executive functions were contrasted with the relations between the WJ III Comprehension-Knowledge cluster and the same measures of executive functions, results called into question the distinction between process and content but suggested that there is a common ability or common abilities underlying performance across all of the WJ III and D-KEFS measures.

Keywords: *Woodcock-Johnson; cognitive abilities; executive functions*

Authors' Note: The Woodcock-Muñoz Foundation provided financial support for the study. The opinions expressed in this article do not necessarily reflect those of the Woodcock-Muñoz Foundation. We are thankful for the University of Memphis Campus School, the University of Memphis Psychology Department, and Middle Tennessee State University Psychology Department for allowing access to participants. Portions of this research were presented at the 2005 annual meeting of the American Psychological Association. Please address correspondence to Randy G. Floyd, University of Memphis, 202 Psychology Building, Memphis, TN 38152; e-mail: rgfloyd@memphis.edu.

The Woodcock-Johnson III (WJ III) Tests of Cognitive Abilities (COG; Woodcock, McGrew, & Mather, 2001) was developed to provide reliable and valid measures of a number of important cognitive abilities for individuals ranging from preschool-aged children to persons in late adulthood. The WJ III COG includes 20 tests and yields 21 cluster scores that result from performance on 2 or more tests. Floyd, Shaver, and McGrew (2003) provided a comprehensive review of the extant validity evidence supporting the use and interpretation of these cluster scores. Based on guidelines provided in the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999), they concluded that the cluster scores representing general intelligence (i.e., global ability composites) and the cluster scores representing seven of the broad abilities described in the Cattell-Horn-Carroll (CHC) theory (i.e., the CHC factor clusters) were supported by a well-developed network of validity evidence. Floyd and colleagues (2003) concluded that another set of cluster scores, called *clinical clusters*, were supported by far fewer sources of validity evidence. These clinical clusters are new to the WJ III COG, and 4 were designed to provide measures of working memory, cognitive speed, attention, and executive processes.

The Working Memory cluster was developed to measure the phonological and central executive components of the memory management system called working memory (Baddeley, 1986, 2001). The cluster was also designed to operationalize the CHC narrow ability Working Memory, which is subsumed by the broad ability Short-Term Memory (Flanagan, McGrew, & Ortiz, 2000; McGrew & Woodcock, 2001).

The Cognitive Fluency cluster was developed to measure the speed at which individuals retrieve information from memory, produce words, and make decisions about conceptual similarities. The design of the Cognitive Fluency cluster is consistent with Carroll's (1993) distinction between factors representing level and those representing rate. Tests of level are most frequently scaled so that items become more difficult as examinees progress through them. However, tests of rate focus on speed of performance. Such tests are constructed so that most examinees could complete all items correctly or receive the maximum score if provided enough time. As such, the Cognitive Fluency cluster captures the rate at which examinees perform somewhat complex cognitive tasks.

The Broad Attention and Executive Processes clusters were developed to measure abilities related to executive functions. Executive functions are frequently described as the mental operations that promote the organization of thought and behavior. These operations include organization, mental flexibility, self-directed speech, planning, and problem solving (Lezak, 1995). The Broad Attention cluster was designed to provide a general measure of attention by tapping into qualitatively different aspects of the construct. Based on component models of attention (e.g., Mirsky, 1996; Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991), tests composing this cluster were selected to measure attentional capacity, divided attention, selective attention, and sustained attention (Schrank, Flanagan, Woodcock, & Mascolo, 2002). The Executive Processes cluster was developed to measure the core cognitive

processes associated with executive functions, such as response inhibition, cognitive flexibility, and planning.

Purpose of the Study

The primary purpose of this study is to examine the relations between scores from the four WJ III clinical clusters and measures designed to measure aspects of executive functions. At present, no study has surfaced that has examined these relations. Consistent with studies examining criterion-related validity or convergent validity, this study describes the statistical associations between the WJ III Working Memory, Cognitive Fluency, Broad Attention, and Executive Processes cluster scores and select measures from the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001) in a sample of children and a sample of adults who were undergraduate and graduate students. The D-KEFS includes the only conormed executive function subtests with a nationally representative standardization sample for individuals within this broad age range, so it provides some of the most ideal measures of executive functions available as criterion variables.

A secondary purpose of this study is more conceptual and theoretical. That purpose is to contrast (a) the probable convergent relations between the process-oriented WJ III clinical clusters and the measures of executive functions from the D-KEFS and (b) the relations between a content-oriented measure from the WJ III and the same D-KEFS measures. The content-oriented measure is the WJ III Comprehension-Knowledge cluster, which measures vocabulary and cultural knowledge. According to Carroll (1993), this contrast between process and content is consistent with Cattell's (1943) Fluid Intelligence and Crystallized Intelligence and Hebb's (1942) Intelligence A and Intelligence B. Fluid intelligence and Intelligence A represent process-oriented abilities associated with higher order mental processes, and crystallized intelligence and Intelligence B represent content-oriented abilities stemming from exposure to culture and education. Likewise, Luria (1966) distinguished between mental processing and acquired knowledge. Although these conceptions are dated, they are still influential among many engaged in intelligence, cognitive ability, and neuropsychological assessment (e.g., Kaufman & Kaufman, 2004). Based on these conceptions, it could be expected that the WJ III Comprehension-Knowledge cluster, which many view as operationalizing Crystallized Intelligence, would demonstrate correlations with the measures of executive functions that were lower in magnitude than the corresponding correlations between those clusters measuring processes (i.e., the WJ III clinical clusters).

On the other hand, following the body of research focusing on general intelligence (e.g., Carroll, 1993; Jensen, 1998; Spearman, 1927), it could be expected that the WJ III Comprehension-Knowledge cluster would demonstrate as high or higher correlations with the D-KEFS measures because measures of Crystallized Intelligence tend to be highly saturated by general intelligence. Thus, they tend to demonstrate relatively high, positive correlations with most other measures of cognitive abilities. If the WJ III clinical clusters are less saturated by general intelligence than the WJ III Comprehension-Knowledge cluster, they would be expected to demonstrate lower correlations with the

D-KEFS measures. This study should provide insights into the accuracy of these conceptions of the similarities and differences between cognitive ability measures.

Method

Participants

Child sample. Participants included 92 children (48 girls and 44 boys) ranging from 8 to 18 years of age (M = 133.38 months, SD = 29.77 months). Approximately 84% of the participants were White (n = 76), 12% were African American or Black (n = 11), and 3% were Asian/Pacific Islanders (n = 3). Using father's education level as an index of socioeconomic status (SES), 1% of fathers did not complete high school (n = 1), 19% completed high school (n = 17), 36% attended some college (n = 31), and 44% obtained a college degree (n = 38). The primary language of all participants was English.

Children were recruited from two school districts in Tennessee. Letters to parents of children enrolled in Grades 3 through 12 were distributed by classroom teachers. Parents were asked to return the informed consent to their child's school or to mail it to the researchers via a postage-paid envelope. In addition, informed consent was obtained from all child participants prior to testing. Although all children attended general education classrooms, some parents reported that their children had been previously diagnosed with educational or psychological conditions. At the completion of testing, children were provided a gift card to a department store.

Adult sample. Participants included 100 adults (73 women and 26 men)[1] ranging from 18 to 51 years (M = 285.05 months, SD = 77.50 months). Approximately 69% of the sample were White (n = 69), 24% were African American or Black (n = 24), and 3% were Asian/Pacific Islanders (n = 3). Adult participants included students from two universities recruited through posted invitations to participate in research. Some participants fulfilled course requirements to participate in research, and some received points toward their course grades for participating. There were 31% freshmen (n = 31), 15% sophomores (n = 15), 16% juniors (n = 16), 24% seniors (n = 24), and 10% 1st-year graduate students in psychology (n = 10). Using father's education level as an index of SES, 4% of fathers did not complete high school (n = 4), 31% completed high school (n = 31), 23% attended some college (n = 23), and 35% obtained a college degree (n = 35).[2] The primary language of all participants was English.

Measures

WJ III COG. The WJ III COG is a battery of 20 tests of cognitive abilities that produce standard scores with a mean of 100 and a standard deviation of 15. Independent reviews of the WJ III COG have lauded its development, standardization, and validation process (Cizek, 2003; Sandoval, 2003).

Working Memory results from performance on the Numbers Reversed and Auditory Working Memory tests. Broad Attention results from performance on the Numbers

Reversed, Auditory Working Memory, Auditory Attention, and Pair Cancellation tests. Cognitive Fluency results from performance on the Retrieval Fluency, Decision Speed, and Rapid Picture Naming tests. Executive Processes results from performance on the Concept Formation, Planning, and Pair Cancellation tests. Comprehension-Knowledge results from performance on the Verbal Comprehension and General Information tests. All median reliability coefficients for these clusters are .90 or higher for ages 8 to 18 and for ages 18 to 59 years (McGrew & Woodcock, 2001).

The D-KEFS. The D-KEFS is a battery designed to measure aspects of executive functions. Its tests produce standard scores with a mean of 10 and a standard deviation of 3. Recent reviews have noted its large and nationally representative standardization sample, its sound development, and the innovative interpretive approach it promotes (Dugbartey, 2003; Homack, Lee, & Riccio, 2005; Ramsden, 2003). The battery is composed of 10 tests: Trail Making, Verbal Fluency, Design Fluency, Color/Word Interference, Sorting, 20 Questions, Word Context, Tower, and Proverb. The Trail Making, Verbal Fluency, and Design Fluency tests were omitted from this study because of low reliability, and the Proverb test was not administered for this study because it cannot be administered to all children in the targeted age range. Scores from specific conditions of each remaining D-KEFS test are noted below.

The Color/Word Interference Test measures verbal production and inhibition. Condition 3: Inhibition requires the examinee to name quickly the colors in which words are printed while inhibiting the more automatic task of reading the words. Condition 4: Inhibition/Switching requires the examinee to switch between naming the dissonant ink colors and reading the words. Across ages 8 to 89, Inhibition demonstrated a test-retest reliability coefficient[3] of .75 and Inhibition/Switching a test-retest reliability coefficient of .65 (Delis et al., 2001).

The Sorting Test measures the ability to form categories and infer concepts, oral expression ability, and flexibility of thinking. During the free sorting task, the examinee must sort a series of six cards into two different categories and describe the rule used to generate the categories for each set of cards. The Free Sorting Description represents the accuracy and quality of the examinee's descriptions of the categorization rules. During the sort recognition task, the examiner sorts the cards into categories, and the examinee must describe the rule used to generate the categories. The Sort Recognition Description also represents the accuracy and quality of the examinee's descriptions of the categorization rules. The alternate-form reliability coefficient[4] was .39 for the Free Sorting Description and .72 for the Sort Recognition Description. Across ages 8 to 89, the Free Sorting Description demonstrated a test-retest reliability coefficient of .75, and the Sort Recognition Description demonstrated a test-retest reliability coefficient of .65 (Delis et al., 2001).

The 20 Questions Test provides a measure of the ability to generate and test hypotheses. Across four trials, the examinee is required to ask the fewest yes-or-no questions as possible to identify a target object from a set of 30 common objects presented in a visual array. The Total Weighted Achievement score represents the total number of questions asked by the examinee to arrive at the target objects across trials, after adjustment for

guessing. Across ages 8 to 89, the median Spearman-Brown corrected split-half reliability coefficient across the four trials of the test was .82, and the test-retest reliability coefficient was .24. The alternate-form reliability coefficient was .37.

The Word Context Test measures deductive reasoning, hypothesis testing, and flexibility of thinking. For each of 10 items, the examinee is required to deduce the meanings of a made-up word using clues stemming from use of the word in five sentences. The Total Consecutively Correct score represents the efficiency of providing correct responses to clue sentences and the continued provision of the correct response to the remaining clue sentences for that item. Across ages 8 to 89, the median Spearman-Brown corrected split-half reliability coefficient across the 10 items was .63, and the test-retest reliability coefficient was .70.

The Tower Test measures spatial planning, rule learning, and inhibition of impulsive and preservative responding. The examinee is required to move disks of varying size across three pegs to build a designated tower using the fewest number of moves possible. The Total Achievement score represents the number of disk moves required to correctly reproduce the designated towers. Across ages 8 to 89, the median Spearman-Brown corrected split-half reliability coefficient across the two halves of the test was .61, and the test-retest reliability coefficient was .44.

Procedures

The WJ III COG and the D-KEFS were administered in a counterbalanced order by trained graduate students in psychology. Within each test battery, tests were administered in the standard order. Testing was typically conducted during two 2-hour sessions. Standardized scores based on age norms were obtained by using the scoring software associated with each test battery.

All graduate students had completed at least one graduate course focusing on administering, scoring, and interpreting psychological and educational tests. All students participated in training sessions focusing on the two tests batteries and were required to submit evidence of their competency.

Results

Means and standard deviations for variables from each sample are presented in Table 1. Based on a review of standard deviations and their variation from the population standard deviations, all correlations were corrected for restriction or expansion of range (Cohen, Cohen, West, & Aiken, 2003; Gulliksen, 1987).

Child Sample

Uncorrected and corrected Pearson product–moment correlations from the child sample data are presented in Table 2, but only the corrected correlations are discussed. Review of corrected correlations revealed a pattern of significant, positive relations between the WJ III clinical clusters and the D-KEFS measures. For instance, all WJ III

Table 1
Descriptive Statistics for the WJ III Clinical Clusters, the WJ III Comprehension-Knowledge Cluster, and the Select D-KEFS Measures for the Child Sample (n = 90) and the Adult Sample (n = 100)

	Sample			
	Child		Adult	
Measure	*M*	*SD*	*M*	*SD*
WJ III Working Memory	109.3	14.9	107.2	12.5
WJ III Cognitive Fluency	102.3	15.2	106.3	15.7
WJ III Broad Attention	110.8	14.8	107.9	12.1
WJ III Executive Processes	109.2	12.1	105.1	12.7
WJ III Comprehension-Knowledge	105.9	13.6	102.1	12.7
D-KEFS Color/Word: Inhibition	10.3	3.3	11.4	2.4
D-KEFS Color/Word: Inhibition/Switching	10.2	3.4	11.2	2.3
D-KEFS Sorting: Free Sorting Description	10.5	2.5	11.0	2.5
D-KEFS Sorting: Sort Recognition Description	10.3	2.9	10.0	3.4
D-KEFS 20 Questions: Total Weighted Achievement	10.4	3.2	11.1	2.2
D-KEFS Word Context: Total Consecutively Correct	10.6	2.9	11.0	2.6
D-KEFS Tower: Total Achievement	10.4	2.6	10.7	2.4

Note: WJ III = Woodcock-Johnson III Tests of Cognitive Abilities; D-KFES = Delis-Kaplan Executive Function System.

clinical clusters were significantly and positively correlated with Color/Word: Inhibition, Color/Word: Inhibition/Switching, Sorting: Free Sorting Description, Sorting: Sort Recognition Description, 20 Questions: Total Weighted Achievement, and Word Context: Total Consecutively Correct. In contrast, none of the WJ III clinical clusters were significantly correlated with Tower: Total Achievement. Noteworthy findings for each clinical cluster are noted below.

Whereas most of the significant correlations between Working Memory and the D-KEFS measures were in the weak range, its relations with Word Context: Total Consecutively Correct and 20 Questions: Total Weighted Achievement were in the moderate range.[5] Similarly, most of the significant correlations between Cognitive Fluency and the D-KEFS measures were in the weak range. Only its correlation with Color/Word: Inhibition was in the moderate range. Consistent with the findings from the Working Memory cluster, Broad Attention demonstrated moderate correlations with Word Context: Total Consecutively Correct and 20 Questions: Total Weighted Achievement. In addition, its correlation with Color/Word: Inhibition was also in the moderate range. Other significant correlations were in the weak range. Both Broad Attention and Working Memory demonstrated their strongest relations with 20 Questions: Total Weighted Achievement. In contrast to the other three clinical clusters, all significant correlations between the Executive Processes cluster and the D-KEFS measures were in the moderate range. Its highest correlation was with Sorting: Sort Recognition Description.

Table 2
Correlations Between the WJ III Clusters and Select D-KEFS Measures for the Child Sample

	WJ III Clinical Cluster									
	Working Memory		Cognitive Fluency		Broad Attention		Executive Processes		Gc Cluster	
D-KEFS Measure	U	C	U	C	U	C	U	C	U	C
Color/Word: Inhibition	.37**	.37**	.50**	.49**	.46**	.46**	.44**	.52**	.40**	.43**
Color/Word: Inhibition/ Switching	.30**	.30**	.35**	.34**	.37**	.38**	.35**	.42**	.38**	.41**
Sorting: Free Sorting Description	.29**	.29**	.31**	.31**	.32**	.33**	.35**	.42**	.44**	.48**
Sorting: Sort Recognition Description	.30**	.31**	.40**	.39**	.38**	.39**	.50**	.58**	.48**	.51**
20 Questions: Total Weighted Achievement	.46**	.47**	.24*	.24*	.51**	.51**	.44**	.52**	.34**	.37**
Word Context: Total Consecutively Correct	.42**	.43**	.39**	.39**	.45**	.46**	.48**	.56**	.61**	.65**
Tower: Total Achievement	.14	.15	–.00	.00	.10	.10	.12	.15	.06	.06

Note: WJ III = Woodcock-Johnson III Tests of Cognitive Abilities; D-KFES = Delis-Kaplan Executive Function System; Gc = Comprehension-Knowledge; U = uncorrected correlation; C = corrected correlation. $n = 92$.
*$p < .05$, two-tailed. **$p < .01$, two-tailed.

Like the WJ III clinical clusters, the WJ III Comprehension-Knowledge cluster was positively and significantly correlated with all D-KEFS measures with the exception of Tower: Total Achievement. All significant correlations, except the correlation between Comprehension-Knowledge and 20 Questions: Total Weighted Achievement, were in the moderate range. Its highest correlation was with Word Context: Total Consecutively Correct. In fact, this correlation (.65) was the highest of any presented in Table 2. To compare the magnitude of the correlations between the Comprehension-Knowledge cluster and each D-KEFS measure to the correlations between each clinical cluster and each D-KEFS measure, all uncorrected correlations were converted to z scores and then tested for significant differences (Cohen et al., 2003). Only 4 of the 28 corrected correlations (14%) between a clinical cluster and a D-KEFS measure were significantly different than their counterpart correlations between the Comprehension-Knowledge cluster and the same D-KEFS measures ($p < .05$). In all instances, the correlations between the

Table 3
Correlations Between the WJ III Clusters and Select D-KEFS Measures for the Adult Sample

	WJ III Clinical Cluster									
	Working Memory		Cognitive Fluency		Broad Attention		Executive Processes		Gc Cluster	
D-KEFS Measure	U	C	U	C	U	C	U	C	U	C
Color/Word: Inhibition	.16	.19	.34**	.33**	.26**	.32**	.34**	.39**	.09	.11
Color/Word: Inhibition/ Switching	.21*	.25**	.46**	.44**	.31**	.37**	.25**	.29**	.17	.20*
Sorting: Free Sorting Description	.35**	.41**	.17	.16	.24**	.29**	.25**	.29**	.48**	.55**
Sorting: Sort Recognition Description	.40**	.46**	.07	.07	.34**	.41**	.38**	.44**	.44**	.50**
20 Questions: Total Weighted Achievement	.10	.12	.07	.07	.08	.10	.17	.20*	.00	.00
Word Context: Total Consecutively Correct	.56**	.63**	.18	.18	.48**	.56**	.39**	.45**	.51**	.58**
Tower: Total Achievement	.16	.19	–.05	–.05	.15	.18	.21*	.25**	.26**	.31**

Note: WJ III = Woodcock-Johnson III Tests of Cognitive Abilities; D-KFES = Delis-Kaplan Executive Function System; Gc = Comprehension-Knowledge; U = uncorrected correlation; C = corrected correlation. $n = 100$.
*$p < .05$, two-tailed. **$p < .01$, two-tailed.

Comprehension-Knowledge cluster and the D-KEFS measures were higher in magnitude than those for the clinical clusters. Comprehension-Knowledge yielded a significantly higher correlation with Sorting: Sort Recognition Description than Working Memory and a significantly higher correlation with Word Context: Total Consecutively Correct than did Working Memory, Cognitive Fluency, and Broad Attention.

Adult Sample

Uncorrected and corrected Pearson product–moment correlations from the adult sample data are presented in Table 3, but only the corrected correlations are discussed. Review of corrected correlations revealed a somewhat different pattern of results than those from the child sample. Whereas there was a pattern of significant, positive relations between the clinical clusters and all D-KEFS measures except for Tower: Total

Achievement in the child sample, in the adult sample, all clinical clusters were significantly correlated with only Color/Word: Inhibition/Switching. In addition, only Executive Processes was significantly correlated with 20 Questions: Total Weighted Achievement. Furthermore, Executive Processes was significantly correlated with Tower: Total Achievement in the adult sample, whereas no clinical cluster demonstrated such relations in the child sample. The patterns of relations for each clinical cluster are noted below.

Working Memory was significantly correlated with four of seven D-KEFS measures: Color/Word: Inhibition/Switching, Sorting: Free Sorting Description, Sorting: Sort Recognition Description, and Word Context: Total Consecutively Correct. Its correlations with Sorting: Free Sort Description, Sorting: Sort Recognition Description, and Word Context: Total Consecutively Correct were in the moderate range, and its highest correlation was with Word Context: Total Consecutively Correct. Cognitive Fluency was significantly correlated with only two D-KEFS measures: Color/Word: Inhibition and Color/Word: Inhibition/Switching. Only its correlation with Color/Word: Inhibition/Switching was in the moderate range. Broad Attention was significantly correlated with five D-KEFS measures: Color/Word: Inhibition, Color/Word: Inhibition/Switching, Sorting: Free Sorting Description, Sorting: Sort Recognition Description, and Word Context: Total Consecutively Correct. Consistent with the findings from the Working Memory cluster, Broad Attention demonstrated correlations in the moderate range with Sorting: Sort Recognition Description and Word Context: Total Consecutively Correct. Executive Processes was significantly correlated with all of the D-KEFS measures, including Tower: Total Achievement. Its correlations with Sorting: Sort Recognition Description and Word Context: Total Consecutively Correct were both in the moderate range, and its highest correlation was with Word Context: Total Consecutively Correct.

Correlations between the WJ III Comprehension-Knowledge cluster and five D-KEFS measures were statistically significant and positive. Its correlations with Sorting: Free Sorting Description, Sorting: Sort Recognition Description, and Word Context: Total Consecutively Correct were in the moderate range, and its highest correlation was with Word Context: Total Consecutively Correct. After all correlations were converted to z scores and then tested for significant differences (Cohen et al., 2003), results showed that only 9 of the 28 corrected correlations (32%) between a clinical cluster and a D-KEFS measure were significantly different than their counterpart correlations between Comprehension-Knowledge and the same D-KEFS measures ($p < .05$). Comprehension-Knowledge yielded a significantly higher correlation with Sorting: Free Description than did Cognitive Fluency, Broad Attention, and Executive Processes. It also demonstrated significantly higher correlations with Sorting: Sort Recognition Description, Word Context: Total Consecutively Correct, and Tower: Total Achievement than did Cognitive Fluency. Only three clinical clusters demonstrated significantly higher correlations with a D-KEFS measure than Comprehension-Knowledge did. Cognitive Fluency yielded a significantly higher correlation with Color/Word: Inhibition/Switching, and Broad Attention and Executive Processes yielded significantly higher correlations with Color/Word: Inhibition.

Discussion

Clinical Clusters

Across both samples, all WJ III clinical clusters demonstrated evidence of statistically significant, positive, and moderate relations with at least some measures of executive functions. These relations were more consistent and generally stronger with the sample of children. Consistent with the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), these results provide some evidence based on external relations that supports the validity of these measures from the WJ III (Woodcock et al., 2001). Although is it important that additional validity evidence surfaces to support the interpretation and use of these scores, the results of this study add to the existing validity evidence supporting these clusters that was summarized by Floyd et al. (2003).

Working Memory and Broad Attention. The external relations between the Working Memory cluster and the D-KEFS scores seem to support its goal of measuring the ability to hold information in immediate awareness while performing a mental operation on that information. Patterns of correlations from both the child and adult samples indicate that it correlated moderately and significantly with other measures that require this ability. It is notable that Working Memory demonstrated some of its strongest correlations with Word Context: Total Consecutively Correct across both samples. For the child sample, its correlation with 20 Questions: Total Weighted Achievement is its highest. Both of these D-KEFS tasks tap heavily the ability of examinees to hold information in mind. For Word Context, examinees must hold at least the meaning conveyed by up to five sentences in mind to deduce the meanings of made-up words. For 20 Questions, examinees must hold in mind the questions they have already asked and what objects were eliminated while they generate other questions to identify the target object.

Because the Broad Attention cluster results, at least partially, from performance on the same two tests as the Working Memory cluster does, it is no surprise that the Broad Attention and Working Memory clusters demonstrated similar patterns of relations with the D-KEFS measures across both samples with only a few exceptions. Those exceptions included somewhat stronger relations between Broad Attention and the D-KEFS measures that were timed and in which points are awarded for speedy performance, Color/Word: Inhibition and Color/Word: Inhibition/Switching. Overall, in terms of external relations, perhaps little is gained from the use of the Broad Attention cluster above that already provided by the Working Memory cluster. If anything is gained, it may be the emphasis on focused attention facilitating quick decisions and responses.

Cognitive Fluency. The external relations between the Cognitive Fluency cluster and the D-KEFS scores largely support its goal of measuring the speed and ease by which an individual performs cognitive operations. With the adult sample, it was significantly and moderately correlated with only the two D-KEFS measures that are timed, Color/Word: Inhibition and Color/Word: Inhibition/Switching. With the child sample, Cognitive Fluency was significantly correlated with the vast majority of the

D-KEFS measures, but it demonstrated its highest correlation with Color/Word: Inhibition. This pattern of external relations between the Cognitive Fluency cluster and the D-KEFS measures is not surprising because all tests contributing to the Cognitive Fluency cluster are timed and points are awarded for speedy performance. In addition, two of the three tests contributing to the Cognitive Fluency cluster require rapid production of words via speech (i.e., Retrieval Fluency and Rapid Picture Naming), and both Color/Word: Inhibition and Color/Word: Inhibition/Switching do as well.

Executive Processes. The Executive Processes cluster was designed to measure the core cognitive processes associated with executive functions such as response inhibition, cognitive flexibility, and planning, and its relations with the D-KEFS scores largely support its design goals. For instance, across all correlations between the WJ III clinical clusters and the D-KEFS measures, Executive Processes had the highest arithmetic average (M = 0.45 from the child sample and M = 0.33 from the adult sample). Executive Processes also was the only clinical cluster that demonstrated significant relations with all seven D-KEFS measures in the adult sample. In fact, its statistically significant (but still somewhat weak) correlations with Tower: Total Achievement provided support for its measurement of executive function beyond that of any other WJ III clinical cluster. Consistent with the design goals of the Executive Processes cluster, these findings indicate that, when compared to the other WJ III clinical clusters, it demonstrated the strongest relations with measures of executive functions.

Characteristics of samples. Although differences in the magnitude of the correlations between the two samples may be specific to our samples, they may reflect actual differences in the abilities tapped or the strategies used at different ages. Because there is some evidence of invariance in the abilities measured by most of the WJ III tests included in this study (Taub & McGrew, 2004) and because the requirements of the D-KEFS tests seem to allow for the greatest variety of strategy use (Delis et al., 2001), it is probable that these differences would be reflected most on the D-KEFS tests. For example, for children, 20 Questions: Total Weighted Achievement may tap into great amounts of working memory resources (and contribute to the most variance in scores) because children differ more in the ability to hold previous questions in mind while developing alternate questions. In contrast, most adults enrolled in higher education are likely to have adequately large working memory space and sufficient working memory resources to handle these task requirements. On the other hand, adults may choose to employ strategies calling on working memory resources to complete the tests such as Sorting and Word Content, whereas children do not. Additional research is needed to examine the reason for these differences in correlations across such age groups.

Content and Process

Comprehension-Knowledge is a measure of vocabulary and cultural knowledge. Based on the conception of content-oriented and process-oriented measures, it was

predicted that Comprehension-Knowledge would demonstrate consistently and notably lower correlations with the measures of executive functions than would the WJ III clinical clusters. The results of this study reveal that this prediction was largely incorrect. Only 3 of 56 corrected correlations between the WJ III clinical clusters and the D-KEFS measures (5%) across the two samples were significantly higher than the respective correlations with the Comprehension-Knowledge cluster. Furthermore, Comprehension-Knowledge demonstrated significantly higher correlations than the WJ III clinical clusters on 10 occasions (18%). There were no significant differences between these correlations for more than three quarters of the comparisons. Such results call into question the accuracy of the distinction between content and process. Perhaps abilities may be better conceptualized as lying on a continuum between (a) those that are specialized and acquired through direct instruction in formal educational contexts and (b) those that are more general and acquired through observation and trial-and-error throughout development (see Flanagan, Ortiz, Alfonso, & Mascolo, 2002, pp. 48-50). The almost uniform pattern of positive and significant correlations across measures suggests that there is a common ability (e.g., general intelligence) or common abilities underlying performance on all of the WJ III and D-KEFS measures. On the other hand, some may argue that it is not a cognitive ability but actually executive functions that underlie performance on these measures.

Limitations

These results offer insight into the relations between the measures from the WJ III and those from the D-KEFS, but there are at least two limitations that limit their applicability. First, participants were drawn from two localized populations in two cities in Tennessee. As a result, across samples, the means scores for the vast majority of the measures were at least slightly above the population mean. Similarly, many variables from these samples demonstrated restriction or expansion of range. This limitation focusing on the generalizability of these findings must be considered in the context that both samples are relatively large (more than 90) for a study such as this one and that corrected correlations were presented to provide values closer to population estimates. Second, by focusing on measured variables and correlations between them, this study did not examine the latent abilities tapped by these measured variables. It is likely that factor analytic techniques that consider patterns of relations at more than one order or stratum will improve the understanding of similarities and differences between measures of cognitive abilities and measures of executive functions.

Notes

1. Gender was not reported for 1 participant.

2. Father's education level was not reported for 7 participants (7%).

3. Delis, Kaplan, and Kramer (2001) reported that the test-retest interval "ranged from 9 to 74 days, with an average time between administrations of 25 ± 12.8 days" (p. 18).

4. Delis et al. (2001) reported that alternate-form reliability estimates stemmed from 286 examinees ages 16 to 89 who completed the standard and alternate forms of selected tests in a counterbalanced order.

5. We recognize that there is no standard rule of thumb for providing nominal labels for *r* values. We drew from the following general labels: *negligible to very weak* (.0 to .2), *weak* (.2 to .4), *moderate* (.4 to .7), *high* (.7 to .9), and *very strong* (.9 to 1.0).

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Baddeley, A. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. (2001). Is working memory still working? *American Psychologist*, *56*, 851-864.

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

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*, 153-193.

Cizek, G. J. (2003). Review of the Woodcock-Johnson III. In B. S. Plake, J. C. Impara, & R. A. Spies, (Eds.), *The fifteenth mental measurements yearbook* (pp. 1020-1024). Lincoln, NE: Buros Institute of Mental Measurements.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system*. San Antonio, TX: Psychological Corporation.

Dugbartey, A. T. (2003). Review of the Delis-Kaplan executive function system. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* (pp. 281-283). Lincoln, NE: Buros Institute of Mental Measurements.

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

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Mascolo, J. T. (2002). *The achievement test desk reference (ATDR): A comprehensive framework for LD determination*. Boston: Allyn & Bacon.

Floyd, R. G., Shaver, R. B., & McGrew, K. S. (2003). Interpretation of the Woodcock-Johnson III Tests of Cognitive Abilities: Acting on evidence. In F. A. Schrank & D. P. Flanagan (Eds.), *WJ III clinical use and interpretation* (pp. 1-46, 403-408). New York: Academic Press.

Gulliksen, H. (1987). *Theory of mental tests*. Hillsdale, NJ: Lawrence Erlbaum.

Hebb, D. O. (1942). The effects of early and late brain injury upon test scores and the nature of normal adult intelligence. *Proceedings of the American Philosophical Society*, *85*, 275-292.

Homack, S., Lee, D., & Riccio, C. A. (2005). Test review: Delis-Kaplan executive function system. *Journal of Clinical and Experimental Neuropsychology*, *27*, 599-609.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children, second edition manual.* Circle Pines, MN: American Guidance Service.

Lezak, M. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.

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

Mirsky, A. F. (1996). Disorders of attention: A neurological perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 71-95). Baltimore: Paul H. Brookes.

Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B., & Kellam, S. G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychology Review*, *2,* 109-145.

Ramsden, P. (2003). Review of the Delis-Kaplan executive function system. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* (pp. 284-286). Lincoln, NE: Buros Institute of Mental Measurements.

Sandoval, J. (2003). Review of the Woodcock-Johnson III. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* (pp. 1024-1028). Lincoln, NE: Buros Institute of Mental Measurements.

Schrank, F. A., Flanagan, D. P., Woodcock, R. W., & Mascolo, J. T. (2002*). Essentials of WJ III cognitive abilities assessment*. New York: Wiley.

Spearman, C. (1927). *The abilities of man: Their nature and measurement.* New York: Macmillan.

Taub, G. E., & McGrew, K. S. (2004). A confirmatory factor analysis of Cattell-Horn-Carroll theory and cross-age invariance of the Woodcock-Johnson Tests of Cognitive Abilities III. *School Psychology Quarterly*, *19*, 72-87.

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