Intern. J. Neuroscience, 117:735–748, 2007

ISSN: 0020-7454 / 1543-5245 online
DOI: 10.1080/00207450600910085

informa
healthcare

EVIDENCE OF CONSTRUCT SIMILARITY IN EXECUTIVE FUNCTIONS AND FLUID REASONING ABILITIES

SCOTT L. DECKER

Georgia State University

SCOTT K. HILL

University of Illinois at Chicago

RAYMOND S. DEAN

Ball State University

Executive Functions and Fluid Reasoning are both considered to be core aspects of intelligence and mediated by frontal lobe functioning. However, both constructs considerably overlap, and the distinction between the two constructs is unclear. For this study, three measures of Executive Functions and three measures of Fluid Reasoning were administered to a group of participants. Significant correlations were found establishing an empirical association between these two constructs. Factor analysis and confirmatory factor analysis also provide evidence for construct similarity. Future research in defining these constructs for measurement purposes and using tests of these constructs in clinical practice is discussed.

Keywords Executive Functions, Fluid Reasoning, Neuropsychology, Cognitive Abilities, Psychometrics, Frontal Lobes

Received 10 April 2006.
Address correspondence to Scott L. Decker, Georgia State University, Counseling and Psychological Services, P.O. Box 3980 Atlanta, GA 30309. E-mail: sdecker@gsu.edu

EXECUTIVE FUNCTIONS AND FLUID REASONING

The term executive functions has been used to describe a variety of cognitive processes involving higher order reasoning, concept formation, shifting of set, novel problem solving, and abstract reasoning (Tarter & Parsons, 1971). Measures such as the Category Test (CT), Trail Making B (TMT-B), and the Wisconsin Card Sorting Test (WCST) have been offered as measures of executive functions (Lezak, 1995); however there is disagreement on what these tests measure. The Wisconsin Card Sorting Test (WCST) (Heaton, Chelune, Talley, Kay, & Curtiss, 1993) has been argued to measure a specific type of attention defined as "shift" (Mirsky, Anthony, Duncan Ahearn, & Kellam, 1991) as well as a result of abstraction ability (Lezak, 1983). Similarly, the Trail Making Test -B has been described as a measure of abstraction (Reitan, 1986) as well as a measure of attention (Lezak, 1987). These discrepancies within the literature are quite widespread and complicate both research in this area and clinical interpretation of these specific tests (O'Donnell, Macgregor, Dabrowski, Oestreicher, & Romero, 1994; Lyon, 1996). In an attempt to offer a general description of all executive function tests, Hayes, Gifford, Ruckstuhl (1996) argue that "... most of these tests involve an unusual circumstances in which subjects are required to perform actions that conflict systematically with immediate and well-established sources of behavioral regulation."

The neuropsychological basis of these measures has been extensively researched and linked to the frontal lobe areas of the brain (Anderson, Bigler, & Blatter, 1995; Arnett, Rao, Bernardin, Grafman, Yetkin, & Lobeck, 1994; Reitan & Wolfson, 1995; Adams & Gilman, 1995). Executive function measures are often directly impacted by injuries to the frontal lobes may also be impacted by injuries to other areas of the brain as well and it is unclear if different executive measures are sensitive to different types of brain injuries (Kolb & Fantie, 1997). In fact, some investigators argue that impaired performance on the Category Test (CT) and the Wisconsin Card Sorting Test (WCST) may represent different types of brain damage or brain regions (Adams & Gilman,1995; Donders & Kirsch, 1991).

The distinction between fluid reasoning and crystallized knowledge, or stores of information in long-term memory, has been a well-researched area within psychology (Cattell, 1963, Horn, 1985). Constructs of fluid and crystallized abilities also have received empirical support through the factor analysis of hundreds of data sets (Carroll, 1993). Fluid reasoning task measure the capacity to reason in novel situations or on novel tasks that cannot be performed automatically using prior knowledge (Stankov, 2003). Thus, tests of fluid reasoning are designed to minimize the use of previously acquired knowledge or previously learned problem-solving procedures on these measures.

Notice, both executive function and fluid reasoning definition share common element in that both involve the application of reasoning strategies to novel or unusual situations. Biological correlates of fluid reasoning have also been of research interests but are less clear than that of executive function measures. Studies have reported selective impairments of fluid abilities in patients with frontal-lobe lesions (Duncan, Burgess, & Emslie, 1995) and Parkinson's disease (Gabrieli, 1996).

The Fluid Reasoning tests of the WJ-R are based on the Fluid-Crystallized theory of cognitive processing, which has evolved from the factor analysis of hundreds of data sets (Woodcock & Mather, 1989). The WJ-R Fluid Reasoning subtests include four measures, although it has since changed in a newer edition of the test. The Concept Formation (CF) sub-test involves categorical reasoning based on principles of formal logic. The Verbal Analogies (VA) sub-test measures the ability to complete phrases with words that indicate appropriate analogies and may be considered a measure of verbal fluid reasoning, as measured in other tests (Roid, 2003). Finally, the Analysis-Synthesis (AS) sub-test measures the ability to analyze the components of an incomplete logic puzzle and to determine and identify the missing components.

Previous research has flexibly interchanged terminology to describe fluid reasoning with executive function terms. The Fluid Reasoning factor of the Woodcock-Johnson has been described as a measure of abstract reasoning (Hessler, 1982). Similarly, the Category Test has been interpreted as being more related to problem solving or "fluid" intelligence than to tasks more reliant on experience or "crystallized" intelligence (Matthews and Reitan, 1963).

Although terminology is often used interchangeably, few empirical studies have been conducted examining these relationships. Telzrow and Harr (1987) examined the relationship between the Wechsler Intelligence Scale for Children-Revised (WISC-R), WJ-R Fluid Reasoning, and the CT. Participants were children with suspected learning disabilities. Results showed that the correlation between Fluid Reasoning Cluster of the WJ-R and VIQ, the FSIQ, and the PIQ were all significant; however, the CT did not correlate significantly with the WJ-R Fluid Reasoning Cluster. In a factor analysis of the CT, WCST, and TMT-B, O'Donnell, Macgregor, Darowski Oestreicher, and Romero (1994) determined that the TMT-B loaded on an attention factor and that the CT and the WCST loaded on what was interpreted as a conceptual factor. If TMT-B is seen as a measurement of executive function, this study is significant in that it suggests the test may measure a different component of executive function than the measures considered.

Another study suggested the CT and the WCST measure different constructs. Perrine (1993), guided by research and theory in cognitive psychology, compared the CT and the WCST on the dimensions of attribute identification and rule learning. Attribute identification involves the extraction of common features in stimuli, which are then categorized and encoded perceptually in memory. Rule learning entails the relation of two or more attributes in a logical rule (and, or, not, etc.). Concept formation was defined as having both attribute identification and rule learning components. Ninety-six neurology patients were administered both the CT and the WCST. In addition, a series of concept formation tests consisting of two rule learning and two logical rule problems were also administered. The results revealed that although the CT and the WCST shared variance, different correlation patterns were found between these measures and measures of concept formation. From these data, Perrine (1993) argued that, although related, the CT and the WCST were not measuring the same construct. Rather, the WCST was correlated with attribute identification task, while the CT was more related to rule learning. Perrine's (1993) study supported past research with the finding of sum 30% shared variance between the two measures.

Given the conceptual similarity but lack of empirical research in investigating executive function and fluid reasoning constructs, the present study investigated the relationship through a variety of methods. First, a correlational method was used to examine the similarity of each test to all other tests. Second, a factor analysis was used to determine the number of underlying constructs among the measures. Finally, a structural equation model (SEM) was used to provide evidence of the number of constructs underlying these measures.

METHOD

Participants

Participants were sixty-seven undergraduate volunteers from a psychology course at a large Mid-Western University. Subjects ranged in age from 19 to 45 years with a mean age of 20.6 (SD = 3.23). There were 25 males and 42 females with a mean of 14.87 years of education. Ninety-two percent of the sample were White, 5% Black and 3% Asian. Eighty-eight percent of the sample was right-handed and 22% was left handed. All subjects denied a history of neurological and/or psychiatric evaluation or treatment.

Measurements

Subjects were administered the CT, TMT-B, WCST and subtests of the WJ-R Fluid Reasoning Cluster in a counter-balanced order. The instruments were counterbalanced in the administration to prevent ordering effects.

All instruments were administered according to standardization instructions for each measure by one of 3 advanced doctoral students in neuropsychology.

All measures were scored in a standard fashion outlined in the administration manual of each test. The total number of errors was computed for both the CT and the WCST while the number of errors and total time was computed for Trails B of the TMT. Consequently, higher scores indicated poor performance for these measures. The Analysis Synthesis, Concept Formation, and Verbal Analogies subtests of the WJ-R were scored according to standard procedures of the WJ-R manual and age corrected scaled scores were computed for each of the subtests and an overall Fluid Reasoning Cluster score. Unlike the other measures, higher scores on the WJ-R subtests indicated better performance, thus, a negative correlation with executive function measures was expected.

Procedure

For the first study, all scores were included in a factor analysis using SPSS 11.0. Principle components analysis was used for factor extraction. In the case multiple factors were found a varimax rotation would be performed.

A structural equation model was created using AMOS 4.0 (Arbuckle & Wothke, 1999). The models were designed to determine the extent executive functions influence fluid reasoning ability. It should be noted the estimated influence of executive functions on fluid reasoning would be the same if the causal arrow were drawn in the opposite direction. Essentially, this model will analyze the extent the underlying latent construct of executive function measures influence the underlying latent construct of fluid reasoning. Since executive functioning has been shown to have different constructs depending upon which test and test variables are examined, two models were examined in this study. The first model was based on the usual use of executive functions test to use the variables that best predict brain damage. The second model used the variables that best represented the latent construct of the tests. This analysis included variables that were not necessarily substantiated in predicting brain damage. In both models, a negative relationship was predicted between executive functions and fluid reasoning, since higher scores on executive functions test indicate greater impairment. In addition, a fairly large relationship was predicted based on definitions of the two constructs and the similar uses of the two constructs in assessments.

Guidelines set forth by Kline (1998) were used to evaluate the confirmatory factor analysis fit statistics. For each of the models, the following indices were examined to evaluate the models fit: 1) chi-square (X^2) goodness-of-fit statistic; 2) X^2 to degrees of freedom ration (X^2/df); 3) Normed Fit Index (NNF): and 4) the Relative Fit Index (RFI). The chi-square goodness-of-fit statistic indicates the overall fit of the data to the model with nonsignificant values indicating a good fit. Similarly, a chi-square and degrees of freedom ratio less than 2.00 indicates a good fit of the data to the model. Both NNF and RFI are used to estimate the relative improvement of a model from a null covariance matrix. The utility of such indexes are in their ease of interpretation. The index can range from 0 to 1 and the closer to 1 the better fit of the data to the model.

Model 1: (Two Uncorrelated Factors). For the first model, Executive Functions and Fluid Reasoning were modeled based on the current description in the literature: two uncorrelated constructs. That is, the underlying construct for executive functions and the underlying construct for fluid reasoning do not share any overlapping variance. This model will provide the best fit if the two constructs are not related. This would also indicate that observed executive function variables (CT, TMT-B, WCST) share no relation with observed fluid reasoning variables (CF, SR, VA, AS). By not specifying any relationship between the two constructs, this model was the most restricted.

Model 2: (Two Correlated Factors). For the second model, two constructs were specified for executive functions and fluid reasoning; however, in this model, the two constructs were specified to be correlated. This model would provide the best fit of the data if tests of executive functions measure a similar construct the tests of fluid reasoning measure.

Model 3: (One-Factor). A third model was tested that specified 1 underlying latent variable that accounts for the majority of the variance for measures of executive functions and fluid reasoning. This model is more parsimonious than model 1 and model 2 since it specifies fewer parameters to be estimated. Like

Table 1. Descriptive statistics (N = 67)

	Descriptive statistics			
	Minimum	Maximum	Mean	Std. deviation
Age in Years	19.00	45.00	20.5522	3.22999
Years Education	14.00	17.00	14.8657	.90278
WCST-Pers	3.00	54.00	11.0448	10.87036
Category	18.00	119.00	60.6418	25.92162
Trails b	18.00	148.00	58.9762	32.95411
Analysis* Synthesis	71.00	129.00	98.9254	12.31938
Concept Formation*	72.00	137.00	99.3433	14.45112
Verbal Analogies*	70.00	135.00	103.6061	13.20475
Valid N (listwise)				

*Score based on normative standard score (100, 15).

model 2, this model is predicted to have the best fit if measures of executive functions overlap with measures of fluid reasoning.

RESULTS

Means and standard deviations for all variables are presented in Table 1. Since scores used in the analysis were based on the standardized residual of partialling out the effect of age, scores are given in z-scores. Table 2 presents the correlation matrix between the fluid reasoning and executive functions variables.

Figures 1, 2, and 3 display the factor structure and the standardized path coefficients for the different models. A summary of fit indices statistics can be found in Table 3.

For the first model (Figure 1), the chi-square indicated a poor fit ($X^2 = 43.332$, df = 14, $p < .000$) as indicated by a statistically significant chi-square. Additionally, the chi-square and degrees of freedom ratio (X^2/df = 3.1) also indicated a poor fit of the data to the model. The modification index for this model suggest a correlation between Executive Functions and Gf would improve the fit of the model (M.I. = 22.93).

The second model (Figure 2) provided a better fit to the data based on the fit statistics. A nonsignificant chi-square ($X^2 = 12.315$, df = 13, $p < .502$). Similarly, the chi-square and degrees of freedom ratio (X^2/df = .947) indicated a good fit of the data to the model. The fit indices also suggested a good fit (NFI = .99, RFI = .98) of the data to the model.

Table 2. Correlations between Executive Function and Fluid Reasoning Measures (N = 67)

	WCST-Per	Category	Trails b	Analysis synthesis	Concept formation	Verbal analogies
WCST-Err Pearson Correlation	1					
Category Pearson Correlation	.390**	1				
Trails b Pearson Correlation	.391**	.394**	1			
Analysis Synthesis Pearson Correlation	−.254*	−.429**	−.340*	1		
Concept Formation Pearson Correlation	−.033	−.343**	−.395**	.332**	1	
Verbal Analogies Pearson Correlation	−.340**	−.406**	−.350**	.350**	.318**	1

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).
WCST-Err (Error Score for the Wisconsin Card Sort Test)

Finally, the third model (Figure 3) was specified to have 1 underlying factor account for the majority of the variance in both executive functions and fluid reasoning measures. A nonsignificant chi-square ($X^2 = 12.948$, df = 14, $p < .531$) indicated a good fit of the data to the model as well as the chi-square/degrees of freedom ratio (X^2/df = .925). The fit indices also indicated a good fit (NFI = .99, RFI = .98).

DISCUSSION

This study provides evidence of association between measures of executive functioning and fluid reasoning. Additionally, this study provides evidence that the common measured element among executive functions tests is similar to the common element measured in fluid reasoning tests. Of all the models tested,

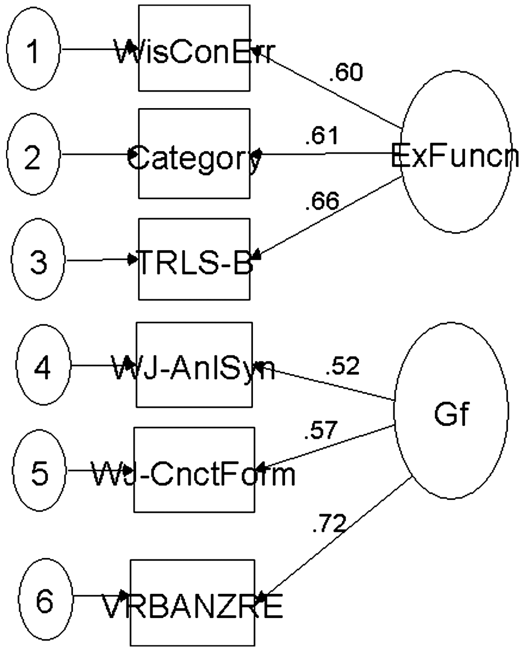

Figure 1. Two-factor Uncorrelated Model.

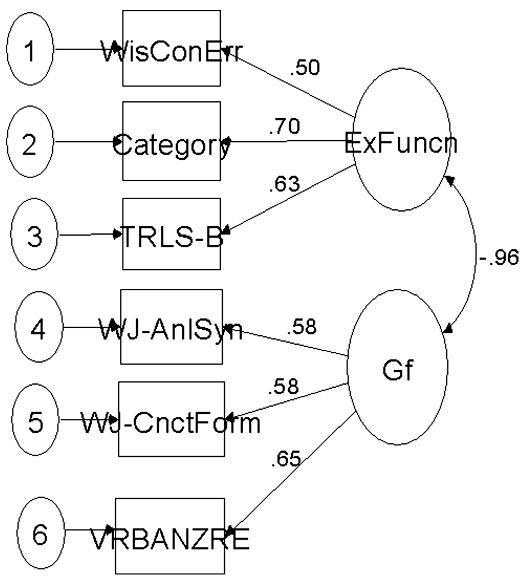

Figure 2. Two-factor correlated model.

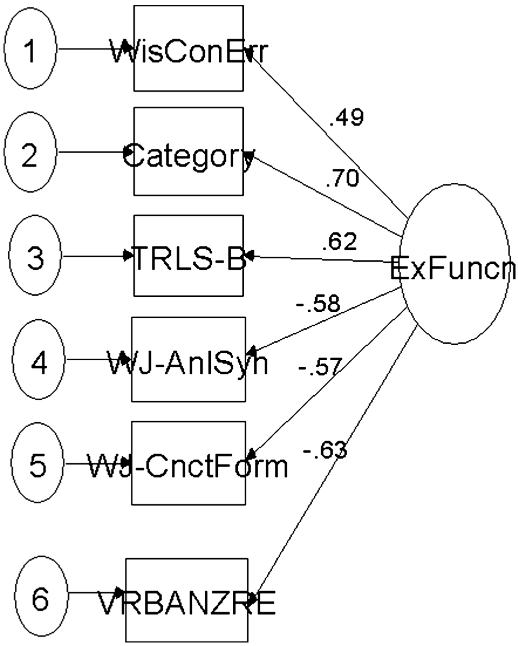

Figure 3. Single Factor model.

only the models that specified a relationship between executive functions and fluid reasoning fit the data. And based on the results, the relationship is fairly strong. Both model 2 and model 3 provided a better fit of the data by specifying a relationship between measures of executive functions and fluid reasoning. Although these results clearly indicate executive functions is related to fluid reasoning, and vice versa, the results do not provide enough information to determine if model 2 is better than model 3. Based on these results, model 2 and model 3 are equivalent and fit the data equally well.

By examining the second model (Figure 2), the structural path indicating the correlation between executive functions on fluid reasoning is.91, which indicates a significant relationship by any standard. Figure 2 also indicates that the structural weights for the executive function variables range from .37 to .61. The lower loading for the WCST may indicate it measures something different than the other two measures of executive functions.

The clinical implication of this study follow from the degree to which clinical studies can be generalized from executive function measures to fluid reasoning measures. It has been demonstrated that certain executive

Table 3. Factor Analysis Results with Executive Function and Fluid Reasoning Measures (N = 67)

	Initial eigenvalues			Extraction sums of squared loadings		
Component	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
		Total variance explained				
1	2.809	46.820	46.820	2.809	46.820	46.820
2	.888	14.803	61.623			
3	.792	13.205	74.828			
4	.581	9.678	84.506			
5	.520	8.668	93.175			
6	.410	6.825	100.000			
		Component matrix (a)				
	Component					
	1					
CATZRE	.764					
VRBANZRE	−.732					
TRLSBZRE	.704					
ANLYSZRE	−.642					
CONCPZRE	−.638					
WCERRZRE	.611					

Extraction Method: Principal Component Analysis.
a 1 components extracted.

function measures are sensitive to detecting frontal lobe damage or frontal lobe dysfunction. Similar in definition, fluid reasoning tasks were compared to several tests of executive functions. Based on the results of the current study, a significant amount of variance of executive functions is shared with fluid reasoning. This suggest the neuropsychological utility of executive function

Table 4. Confirmatory Factor Analysis Model Fit Indicies

Model	X^2	df	p	NFI	TLI	CFI	RMSEA
Two Factor Uncorrelated	40.6	9	.000	.555	.049	.550	.231
Two Factor Correlated	9.8	8	.274	.892	.930	.973	.060
One Factor	10.0	9	.348	.890	.966	.985	.042

Note: NFI = Normed fit index; TLI = Tucker Lewis Index; CFI = Comparative Fit Index; RMSEA = Root-Mean Square Error of Approximation.

measures may generalize or be supplemented by measures of fluid reasoning, but further research is needed for validation.

There are several limitations to this study which may guide future research. First, although the sample size was sufficient to demonstrate an empirical association among the measures, the sample size is too small to make more refined distinctions. Larger sample sizes with greater amounts of variance may detect more subtle variations among the data such as sub-factors involved with attentional processes of shift, focus, and encode (Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991). Second, because many executive function measures have error score metrics, the score distributions may be non-normal and skewed. Future research may consider using alternative metrics as indicators of performance or using non-parametric statistics. Finally, developing a task analysis of each of these tests in order to make predictions of factor analytic results may help clarify the processing demands of each of these tests and elucidate the similarity and differences among the measures.

REFERENCES

Adams, K. M., & Gilman, S. (1995). Correlation of neuropsychological function with cerebral metabolic rate in subdivisions of the frontal lobes of older alcoholic patients measured with fluorodeooxyglucose and positron emission tomography. *Neuropsychology*, *9*(2), 275–280.

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

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.

Denckla, M. A. (1996). A theory and model of executive function. In Lyon, R. G. & Krasnegor, N. A. *Attention, Memory, and Executive Functions*. Baltimore, MD: Paul Brooks. Eslinger, P. (1996). Conceptualization, describing, and measuring components of executive functions.In Lyon, R. G. & Krasnegor, N. A. *Attention, Memory, and Executive Functions*. Baltimore, MD: Paul Brooks

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, *33*, 261–268.

Gabrieli, J. D. E. (1996). Memory systems analyses of mnemonic disorders in aging and age related disease. *Proceedings of the National Academy of Sciences of the United States of America, 93*, 13534–13540.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G.,(1993). *Wisconsin Card Sorting Test manual*. Odessa, FL : Psychological Assessment Resources, Inc.

Hessler, G. L. (1982). Use and interpretation of the Woodcock-Johnson Psycho-Educational Battery. Hingham, MA: Teaching Resources.

Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), Handbook of Intelligence (pp. 267–300). New York: Wiley.

Kolb, B., & Fantie, B. (1997). *Development of the child's brain and behavior*. New York: Plenum Press.

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

Lezak, M. D. (1987). Assessment for rehabilitation planning. In M. Meier, A. Benton, & L. Diller (Eds.), *Neuropsychological rehabilitation*. New York: Guilford Press.

Lezak, M. D. (1983). *Neuropsychological assessment* (2 nd. ed.). New York: Oxford University Press.

Lyon, R. G., & Krasnegor, N. A. (1996). *Attention, Memory, and Executive Function.* Baltimore MD: Paul H. Grookes Publishing Co.

O'Donnell, P. J., Macgregor, L. A, Dabrowski, J. J., Oestreicher, J. M., & Romero, J. J. (1994). Construct validity of neuropsychological tests of conceptual and attentional abilities. *Journal of Clinical Psychology, 50*(4), 596–600.

Mathews, C. G. (1974). Applications of neuropsychological tests methods in mentally retarded subjects. In R. M. Reitan & L. A. Davison (Eds), *Clinical Neuropsychology* (pp. 267–285). Washington, DC: Hemisphere.

McGrew, K. S., Keith, T. Z., Flanagan, D. P., & Vanderwood, M. (1997). Beyond g: The impact of 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.

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*, 109–146.

O'Donnell, J. P., Macgregor, L. A., Dabrowski, J. J., Oestreicher, J. M., & Romero, J. J. (1994). Construct validity of neuropsychological tests of conceptual and attentional abilities. *Journal of Clinical Psychology*, *50*(4), 596–600.

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology*, *1*(4), 349–370.

Perrine, K. (1993). Differential aspects of conceptual processing in the Category Test and Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology*, *15*(4), 461–471.

Reitan, R. M. (1986). Theoretical and methodological bases of the Halstead-Reitan Neuropsychological Battery. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessmentof neuropsychiatric disorders*. New York: Oxford Press.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery.* Tucson, AZ: Neuropsycyhology Press.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales, Fifth Edition, Examiner's Manual*. Itasca, IL: Riverside Publishing Company.

Stankov, L. (2003). Complexity in human intelligence. In R. J. Sternberg, J. Lautrey & T. I. Lubart (Eds.), *Models of intelligence: International perspective* (pp. 27–42). Washington, D.C.: American Psychological Association.

Tarter, R. E., & Parsons, O. A. (1971). Conceptual shifting in chronic alcoholics. *Journal of Abnormal Psychology, 77*, 71–75.

Telzrow, C. F., & Harr, G. A. (1987). Scommon variance among three measures of non-verbal cognitive ability: WISC-R Performance Scale, WJPB-TCA reasoning cluster, and Halstead Category Test. *Journal of School Psychology*, *25*, 93–95.

Waber, D. P. (1977). Sex differences in mental abilities, hemispheric lateralization and rate of physical growth at adolescence. *Devevelopmental Psychology*, *13*, 29–38.

Woodcock, R. W., & Johnson, M. B. (1977). *Woodcock-Johnson Psycho-Educational Battery.* Hingham. MA: Teaching Resources.