Assessment

http://asm.sagepub.com

The Factor Structure and Age-Related Factorial Invariance of the Delis-Kaplan Executive Function System (D-KEFS)

Robert D. Latzman and Kristian E. Markon

Assessment 2010; 17; 172 originally published online Dec 29, 2009;
DOI: 10.1177/1073191109356254

The online version of this article can be found at:
http://asm.sagepub.com/cgi/content/abstract/17/2/172

Published by:

SAGE

http://www.sagepublications.com

Additional services and information for *Assessment* can be found at:

Email Alerts: http://asm.sagepub.com/cgi/alerts

Subscriptions: http://asm.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://asm.sagepub.com/cgi/content/refs/17/2/172

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Assessment
17(2) 172–184
© The Author(s) 2010
Reprints and permission: http://www.sagepub.com/journalsPermissions.nav
DOI: 10.1177/1073191109356254
http://asmnt.sagepub.com

The Factor Structure and Age-Related Factorial Invariance of the Delis–Kaplan Executive Function System (D-KEFS)

Robert D. Latzman[1] and Kristian E. Markon[1]

Abstract

There has been an increased interest in the structure of and relations among executive functions. The present study examined the factor structure as well as age-related factorial invariance of the Delis–Kaplan Executive Function System (D-KEFS), a widely used inventory aimed at assessing executive functions. Analyses were first conducted using data provided in the D-KEFS technical manual and were then replicated in an independent sample of male early adolescents aged 11 to 16 years. Results revealed a three-factor solution best fit the data across groups and samples; measurement properties appeared to be invariant across age groups for certain loadings and variant for others. The three factors were labeled Conceptual Flexibility, Monitoring, and Inhibition. These findings provide better understanding of the measurement properties of the D-KEFS and contribute to the larger literature on the structure of measures intended to assess executive functions.

Keywords

Delis–Kaplan Executive Function System (D-KEFS), executive functions, measurement invariance, factor analysis

Executive function (EF), abilities related to higher order cognitive processes such as judgment, decision making, and planning (Baron, 2003; Lezak, Howieson, & Loring, 2004), are critically important as they represent the potential for both risk as well as protective factors associated with problematic behaviors. As such, there has been an increased interest in the structure of EF and their relationship with other traits and behaviors. The Delis–Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001a), the first inventory of its kind aimed solely at the assessment of EF, provides a unique opportunity for examining the nature of EF in a systematic manner. The D-KEFS includes a variety of measures putatively designed to assess EF, co-normed across a broad age range. Little is known, however, about the factor structure of the D-KEFS, and how this structure relates to what is currently known about the structure of EF more broadly. Here, we examined the factor structure and age-related factorial invariance of the D-KEFS utilizing two separate samples, the D-KEFS standardization sample and a community sample of male early adolescents. Results provide insight into the nature and factor structure of the D-KEFS, and contribute to understanding of EF measurement and structure in general.

Delis–Kaplan Executive Function System

The D-KEFS provides a standardized assessment of EF in individuals between 8 and 89 years. As is evident from reviews published in journals across the field of psychology (Baron, 2004; Homack, Lee, & Riccio, 2005; Skunk, Davis, & Dean, 2006; Swanson, 2005) as well as in data presented in the technical manual of the recently published Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV; Wechsler, 2008), the D-KEFS has quickly begun to be used in a large number of neuropsychology clinics and settings. The D-KEFS consists of various procedures and tasks both in verbal and nonverbal modalities that have demonstrated sensitivity in the detection of frontal lobe (dys)function, functions that have become known as EF (Delis et al., 2001a).

The D-KEFS is the first set of tests designed exclusively for the assessment of executive functioning to be co-normed on a large stratified, representative national sample. The nine tests comprising the D-KEFS are either relatively new or modifications of well-established clinical or experimental tests (Delis et al., 2001a). For example, well-established tests included in the D-KEFS battery include the Color–Word Interference Test, originally developed by Stroop (1935), the Verbal Fluency Test, a modification of the Controlled Oral

[1]University of Iowa, Iowa City, IA, USA

Corresponding Author:
Robert D. Latzman, Department of Psychiatry & Human Behavior, University of Mississippi Medical Center, 2500 North State Street, Jackson, MS 39216, USA
Email: robert-latzman@uiowa.edu

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Word Association Test (COWAT; Benton & Hamsher, 1976), and the Trail Making Test, a modification of the traditional test originally developed by Brown and Partington (1942) and later included in the Halstead–Reitan Neuropsychological Battery (HRNB; Reitan & Wolfson, 1993). Less widely used EF tests included in the battery are the Category Switching Test, which was originally devised by Newcombe (1969), the Word Context Test, and the Sorting Test, a version of the California Category Sorting Test (CCST; Delis, Squire, Bihrle, & Massman, 1992).

Some reviews note adequate validity and reliability data concerning the D-KEFS (Delis, Kramer, Kaplan, & Holdnack, 2004) as well as adequate test–retest reliabilities for most subtests (Homack et al., 2005). Other scholars in the field, however, have expressed concerns that the nine tests included in the battery do not represent a comprehensive assessment of EF (Baron, 2004), and that an insufficient number of reliability and validity studies have been conducted to understand the full utility of the battery (Baron, 2004). To date, moreover, no factor analytic investigation of the structure of the D-KEFS has been reported. As discussed in the technical manual, each D-KEFS test was developed as an individual, stand-alone instrument (Delis, Kaplan, & Kramer, 2001b). Thus, the D-KEFS allows direct comparisons between various tests, making it an excellent candidate for structural studies, both in general, as well as with regard to age-related invariance.

Hypotheses Regarding D-KEFS Factor Structure

One-factor structure. As an inventory designed to assess functioning in a single domain—that of executive functioning—one possibility is that the D-KEFS might exhibit a one-factor structure, with all measures loading strongly on a single factor. Given that previous research on the structure of EF using batteries of widely used executive tasks has often found EF to be somewhat heterogeneous across a wide range of populations (Burgess, 1997; Lehto, 1996; Levin et al., 1996; Lowe & Rabbitt, 1997), such a dominant one-factor structure may be unlikely. However, some previous studies have documented a prominent higher order EF factor (Duncan, Emslie, Williams, Johnson, & Freer, 1996; Duncan, Johnson, Swales, & Freer, 1997; Engle, Kane, & Tuholski, 1999), and it is possible that a higher order EF factor might be observed within a hierarchical framework, with observed measures having loadings from a limited number of EF factors, which in turn have loadings from a single higher order factor.

Two-factor structure: Verbal and nonverbal. Although the authors of the D-KEFS have not reported factor analyses of the inventory, they have speculated on important higher order processes that might affect performance on D-KEFS subtests. In the *D-KEFS Examiner's Manual*, for example, Delis et al. (2001a) note the importance of verbal versus nonverbal processing, and cite evidence from various forms of neuropsychological research that different EF measures differentially involve these two forms of processing. Based on this work, one might speculate that D-KEFS measures would reflect two EF factors, one verbal and the other nonverbal.

Two- and three-factor structures: Inhibition, updating, and shifting. Another possibility is that the D-KEFS, as a comprehensive inventory assessing EF, will exhibit a factor structure similar to that observed in other structural studies of EF. One prominent model of EF that has been gaining increasing attention is a three-factor model, proposed by Miyake et al. (2000), in which the domain of EF is decomposed into three factors: inhibition, updating, and shifting. Inhibition reflects the ability to deliberately control prepotent responses (i.e., responses that are automatic or dominant) and is expressed in performance on measures such as the Stroop task (Miyake et al., 2000). Updating requires monitoring and evaluating the relevance of new information for the task at hand and, if appropriate, updating information in working memory with the new information (Miyake et al., 2000). Finally, the shifting factor described by Miyake et al. requires shifting between multiple operations and performing new operations while being faced with proactive interference, and is expressed in measures such as the Wisconsin Card Sorting Test (WCST).

Recent work on age-related changes in EF factor structure has suggested possible two-factor variants of the Miyake et al. (2000) model. In a recent study of the structure of EF in older adults, for example, Hull, Martin, Beier, Lane, and Hamilton (2008) find that EF tasks reflect two, rather than three, distinct factors, updating and shifting. Similarly, some research in children and adolescents has observed two factors, updating and shifting, without inhibition (Huizinga, Dolan, & van der Molen, 2006; van der Sluis, de Jong, & van der Leij, 2007; see, however, Garon, Bryson, & Smith, 2008, and Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003). The failure of inhibition to emerge as a distinct factor in these younger and older age groups suggests age-related variance in the structure of EF, and raises the possibility that similar two-factor structures might emerge in the D-KEFS in younger and older individuals. This sort of compression in factor structure at younger and older ages would parallel the well-established finding that EF increases with age and then plateaus before slowly declining in older age (Jurado & Rosselli, 2007; Zelazo, Craik, & Booth, 2004).

Goals of the Current Study

The purpose of the current study was to examine the factor structure of the D-KEFS, with a secondary goal of examining the age-related invariance of this structure. Given the

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

stated purpose of the D-KEFS to broadly measure EF, it is important to examine its measurement properties and compare these properties to what is known about existing measures of EF. Also, given the broad age range (8-89 years) of clinical application of the D-KEFS, determining the generality of the measurement model is essential for both researchers as well as clinicians. The present investigation replicated findings across two separate samples. First, exploratory and confirmatory analyses were performed using the standardization sample as reported in the D-KEFS manual. The exploratory findings were then replicated in a second, original community sample of adolescent males aged 11 to 16 years.

Study 1

Method

Sample and Measures. This study used the D-KEFS standardization sample data provided in the technical manual (Delis et al., 2001b). The D-KEFS was standardized on a nationally representative, stratified sample of 1,750 nonclinical children, adolescents, and adults, aged 8 to 89 years. The sample had a minimum of 75 people in each of the age groups that were used to generate age-specific norms. Eight of the D-KEFS tests have been standardized and normed for use with children as young as 8 years: (a) Trail Making Test, which assesses attention, concentration, resistance to distraction, and cognitive flexibility; (b) Verbal Fluency Test, which requires speeded lexical production and automatic lexical access, and reflects efficient lexical organization; (c) Design Fluency Test, an established nonverbal task analogous to verbal fluency; (d) Color–Word Interference Test, which assesses selective or focused attention, the ability to shift from one perceptual set to another as test requirements change, and the ability to inhibit inappropriate responding; (e) Sorting Test, a measure of conceptual flexibility, (f) Twenty Questions Test, which assesses strategic thinking; (g) Word Context Test, which assesses deductive reasoning, hypothesis testing, and flexibility of thinking by requiring participants to discern what is intended by a made-up word based on its use in a series of sentences; (h) Tower Test, which assesses forward planning of a sequence of steps as the participant tries to move a pattern of discs efficiently from a start configuration to a goal configuration to match a target pattern; in addition to a ninth test, (i) the Proverb Test, which is only appropriate for examinees 16 years of age and older and assesses verbal abstraction abilities.

D-KEFS Total Achievement Scores (e.g., Sorting Test Free Sorting Confirmed Correct Sorts and Free Sorting Description Score, Trail Making Test Number–Letter Switching, Color–Word Test Inhibition, Verbal Fluency Test Total Letter Fluency, Verbal Fluency Test Category Switching Total Correct Responses, Tower Test Total Achievement Score) were included in the current analyses. D-KEFS Total Achievement Scores were used because they reflect global achievement scores on these tests (Delis et al., 2001a) and also reflect, in general, traditional measures of EF. The technical manual (Delis et al., 2001b) provided intercorrelations of Total Achievement scaled scores of measures of different D-KEFS tests presented in broad age bands: 8 to 19 years (n = 702), 20 to 49 years (n = 361), and 50 to 89 years (n = 326). As different numbers of individuals completed each combination of measures, the smallest N for any correlation within an age group was used for analyses, to maintain relatively conservative inference about factor models. The Proverbs Test, which is only appropriate for use with individuals aged 16 years and older, was omitted from the 8- to 19-year-old sample but was included for the two older samples by treating the Proverbs Test as missing for the youngest sample. Intercorrelation matrices for each age band were used for analyses.

Analyses. Analyses were conducted in two phases. The goal of the first phase was to characterize the factor structure of the measures in each age group separately, using exploratory approaches. The goal of the second phase was to characterize patterns of age-related factor and measurement invariance, verifying and extending the results of the first phase using confirmatory approaches.

Structural models. Exploratory factor models were fit to correlation matrices of the three age groups separately using maximum-likelihood estimation. Maximum-likelihood estimation has a number of desirable properties (e.g., it is robust to nonnormality when latent factors are allowed to be freely correlated; see Satorra, 2002; Savalei, 2008) and allow for use of maximum-likelihood-based criteria such as Bayesian Information Criterion (BIC). Models comprising between one and eight factors were fit, with loading matrices rotated using quartimin rotation, the default oblique rotation in Mplus (Muthén & Muthén, 1998-2007). An oblique rotation was used because of an a priori expectation that higher order factors would reflect a coherent domain of EF, as suggested by the goals of the D-KEFS (Delis et al., 2001a) and prior research indicating that EF tasks tend to be correlated (e.g., Fisk & Sharp, 2004; Hull et al., 2008; Lehto et al., 2003; Miyake et al., 2000).

Two methods were used to determine the optimal number of factors to model within each age group and use in further analyses: eigenvalue Monte Carlo p values (i.e., parallel analysis; Horn, 1965) and Velicer's minimum average partial (MAP) criterion (Velicer, Eaton, & Fava, 2000). Parallel analysis and MAP have been shown to perform well in identifying the number of factors in an exploratory factor analysis (EFA) model (Hayton, Allen, & Scarpello, 2004; Zwick & Velicer, 1986). Mplus (Muthén & Muthén, 1998-2007) was used for all EFAs.

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Invariance models. To minimize the influence of misspecified factor loading patterns on inference regarding age-related invariance, we used an unrestricted factor analysis (UFA) framework (Jöreskog & Sörbom, 1979) for modeling invariance across age groups. UFA is a form of confirmatory factor analysis in which minimal constraints are placed on the pattern of factor loadings, only as many as are needed to identify the model. Unlike EFA, where all loadings are freely estimated, in UFA, specific assumptions are made about certain loadings. Specifically, for each factor, one measure is identified that is predicted a priori to load on that factor, and to have zero loadings on the other factors; the loadings of all other measures on all the factors are left free. We fixed and freed the loadings of one measure for each factor in this way using the results of the exploratory structural modeling. Mx (Neale, Boker, Xie, & Maes, 2003) allows one to freely estimate parameters, but constrain them to be equal across age groups (i.e., there is one parameter representing the correlation between Factor 1 and Factor 2, which is freely estimated). For each factor, the measure having the highest loading across groups without substantial cross-loadings was used to identify the model.[1]

Invariance models were compared in a series roughly corresponding to standard measurement invariance modeling frameworks. We attempted to make our analyses as close as possible to the standard invariance model comparisons that are designed to evaluate increasingly strict assumptions about the equality of various parameters (e.g., factor loadings, factor intercorrelations). Specifically, we tested a series of model comparisons that represented weak invariance, strict invariance, and partial invariance (Meredith, 1993). The first set of comparisons was intended to evaluate weak factorial invariance across age groups. In this set of comparisons, factor loadings were either fixed or free to vary across age groups, and factor intercorrelations were allowed to vary across groups. The second set of comparisons was intended to evaluate the invariance of factor intercorrelations and strict factorial invariance across age groups. In this step, factor intercorrelations were either fixed or allowed to remain free across age groups. Finally, we evaluated partial factorial invariance models, in which the factor loadings of some measures were constrained to be equal across age groups, and others were allowed to vary. Scores that consistently had prominent loadings on only one factor in EFAs in all groups were fixed to be equal across age groups whereas those scores that did not consistently load on the same factor across groups were freed.

It is important to note that although our model comparisons largely correspond to standard invariance frameworks (Meredith, 1993), there were also a number of differences between the models evaluated here and those frameworks. Most important, given that correlation rather than covariance matrices were modeled, it was not possible to evaluate models of residual observed variance separately from models of factor intercorrelations. Thus, it was not possible to evaluate strict factorial invariance separately from factor intercorrelation invariance. Also, insufficient information about observed measure means was available for invariance modeling, which precluded evaluation of strong factorial invariance. Mx (Neale et al., 2003) was used for invariance analyses because of the unequal number of variables across the three groups.

Results

Structural models. Monte Carlo *p* values and observed eigenvalues are shown in Table 1; factor loadings for the best-fitting exploratory model for each age group are shown in Tables 2, 3, and 4. Results of parallel analyses suggested a three factor solution best fit the data for all groups. MAP analyses suggested a one-factor solution for the two older groups. When MAP did not suggest a one factor solution, as was the case for the 8- to 19-year-old group, it suggested a three-factor solution. EFA models with more than three factors failed to converge in at least one of the age groups, moreover. It was therefore determined that a three-factor solution provided optimal fit to the data across the three groups. To examine the effects of variable selection on these results, as analyses included multiple, correlated scores from many of the subtests, EFAs and parallel analyses were conducted using different combinations of variables, including analyses removing highly correlated measures within the Sorting and Verbal Fluency tasks. Results of parallel analyses with these various scores removed were largely consistent with the results presented in suggesting a three-factor solution.[2]

The first factor is reflected in all three scores from the Sorting Test: Free Sort, Free Sort Description, and Sort Recognition. This was consistent for all age groups. The various scores that make up the Sorting Test which uniformly load on the first factor measure abilities such as initiation of problem solving, concept formation, ability to explain concepts, ability to transfer abstract concepts into action, the ability to inhibit previous responses, both verbally and behaviorally, to engage in flexible thinking and behavior. This factor was therefore labeled Conceptual Flexibility as it requires abilities related to abstraction and concept formation (Greve, Farrell, Besson, & Crouch, 1995).

The second factor was anchored by the two Category Switching scores from the Verbal Fluency tests for all groups. In general, loadings on the remaining two Verbal Fluency scores, Letter Fluency and Category Fluency, were more modest, and tended to decrease with age. This factor was labeled Monitoring, to emphasize the verbal and semantic monitoring demands of the Category Switching tasks (total correct and accuracy) of the Verbal Fluency tests, as

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Table 1. Study 1: Observed Eigenvalues and Monte Carlo *p* Values

	Eigenvalue no.							
	1	2	3	4	5	6	7	8
8-19-year olds								
Observed eigenvalue	4.333	1.724	1.457	1.120	1.029	0.915	0.903	0.801
Monte Carlo *p*	.000	.000	.000	.482	1.000	1.000	1.000	1.000
20-49-year-olds								
Observed eigenvalue	5.243	1.752	1.358	1.056	0.975	0.888	0.840	0.714
Monte Carlo *p*	.000	.000	.000	1.000	1.000	1.000	1.000	1.000
50-89-year-olds								
Observed eigenvalue	5.866	1.514	1.376	1.017	0.967	0.895	0.755	0.635
Monte Carlo *p*	.000	.000	.000	1.000	1.000	1.000	1.000	1.000

Note: The *p* values were calculated by Monte Carlo methods as described in the text.

Table 2. Study 1: Quartimin Rotated Exploratory Factor Model for 8- to 19-Year-Olds

D-KEFS Achievement Tests	Conceptual Flexibility	Monitoring	Inhibition	h^2
Sorting Cond. 1: Free Sort	**.97**	–.00	–.02	.07
Sorting Cond. 2: Free Sort Description	**1.00**	–.02	–.01	.03
Sorting Cond. 3: Sort Recognition	**.53**	.13	.12	.60
Trail Making Test	.07	.10	**.38**	.81
Color–Word Test: Inhibition	–.03	.01	**.82**	.34
Color–Word Test: Inhibition/Switching	.02	–.08	**.69**	.54
Twenty Questions Test	.20	.02	.20	.89
Verbal Fluency: Letter Fluency	.10	**.37**	.25	.71
Verbal Fluency: Cat. Fluency	.06	**.46**	.21	.67
Verbal Fluency: Cat. Switch total	.00	**.92**	–.02	.17
Verbal Fluency: Cat. Switch Accuracy	–.03	**.80**	–.05	.40
Design Fluency	.07	.08	.25	.90
Tower Test: Achievement	–.11	–.01	–.05	.98
Tower Test: Accuracy Ratio	.25	.24	.17	.77
Word Context Test	.24	.19	.18	.81

Note: D-KEFS = Delis–Kaplan Executive Function System; $N = 702$; h^2 = unique variance. Loadings ≥.30 are in given in boldface.

Table 3. Study 1: Quartimin Rotated Exploratory Factor Model for 20- to 49-Year-Olds

D-KEFS Achievement Tests	Conceptual Flexibility	Monitoring	Inhibition	h^2
Sorting Cond. 1: Free Sort	**.99**	–.02	–.02	.04
Sorting Cond. 2: Free Sort Description	**1.00**	.01	–.03	.02
Sorting Cond. 3: Sort Recognition	**.54**	.07	.28	.48
Trail Making Test	.13	.07	**.50**	.65
Color–Word Test: Inhibition	–.08	.03	**.70**	.52
Color–Word Test: Inhibition/Switching	–.03	–.04	**.78**	.45
Twenty Questions Test	.14	.14	.17	.88
Verbal Fluency: Letter Fluency	.11	.23	**.35**	.70
Verbal Fluency: Cat. Fluency	.11	**.39**	.28	.62
Verbal Fluency: Cat. Switch Total	–.01	**.97**	–.03	.10
Verbal Fluency: Cat. Switch Accuracy	–.01	**.84**	.01	.30
Design Fluency	.21	–.02	**.39**	.75
Tower Test: Achievement	.13	–.11	**.31**	.89
Tower Test: Accuracy Ratio	–.09	–.14	.06	.97
Word Context Test	.22	.14	.29	.75
Proverbs Test	.21	.18	.28	.74

Note: D-KEFS = Delis–Kaplan Executive Function System; $N = 361$; h^2 = unique variance. Loadings ≥.30 are in given in boldface.

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Table 4. Study 1: Quartimin Rotated Exploratory Factor Model for 50- to 89-Year-Olds

D-KEFS Achievement Tests	Conceptual Flexibility	Monitoring	Inhibition	h^2
Sorting Cond. 1: Free Sort	**.99**	–.00	–.01	.04
Sorting Cond. 2: Free Sort Description	**.99**	.01	–.01	.02
Sorting Cond. 3: Sort Recognition	**.50**	.03	.26	.52
Trail Making Test	.12	.04	**.57**	.56
Color–Word Test: Inhibition	–.08	–.00	**.65**	.63
Color–Word Test: Inhibition/Switching	–.05	–.06	**.64**	.66
Twenty Questions Test	.21	.09	.11	.89
Verbal Fluency: Letter Fluency	.03	.03	**.61**	.59
Verbal Fluency: Cat. Fluency	–.06	.14	**.60**	.59
Verbal Fluency: Cat. Switch Total	.01	**.97**	–.01	.05
Verbal Fluency: Cat. Switch Accuracy	–.02	**.79**	.01	.38
Design Fluency	.07	.01	**.46**	.74
Tower Test: Achievement	–.03	–.00	**.61**	.65
Tower Test: Accuracy Ratio	–.08	.12	**–.37**	.86
Word Context Test	.26	–.00	**.45**	.60
Proverbs Test	.24	.03	**.40**	.65

Note: D-KEFS = Delis–Kaplan Executive Function System; $N = 26$; h^2 = unique variance. Loadings ≥.30 are in given in boldface.

Table 5. Study 1: Estimated Interfactor Correlations for EFAs and Best-Fitting UFA

Age Group	Conceptual Flexibility and Monitoring	Conceptual Flexibility and Inhibition	Monitoring and Inhibition
EFAs			
8-19-year-olds	.26	.32	.26
20-49-year-olds	.26	.38	.45
50-89-year-olds	.30	.58	.45
Best-fitting UFA	.25	.30	.21

Note: EFA = exploratory factor analysis; UFA = unrestricted factor analysis. Interfactor correlation estimates were calculated separately for each age group in the context of separate EFAs as described in the text. Interfactor correlation estimates for the best-fitting UFA model were calculated in the context of a multigroup UFA.

opposed to the production demands that characterize the other Verbal Fluency subtests (Troyer, Moscovitch, & Winocur, 1997; Nutter-Upham et al., 2008) which have more modest loadings on this factor. The Monitoring factor appears to be closely linked to working memory and requires the ability to engage and disengage appropriate task sets in addition to performing new operations in the face of interference. The abilities tapped by this factor are active rather than passive in that they require conscious manipulation of the content of working memory.

The third factor was anchored by the Inhibition and Inhibition/Switching scores from the Color-Word Test for all groups in addition to having the Trail Making Test and Design Fluency across age groups. For the oldest group, the 50- to 89-year-olds, both scores from the Tower Test, the Word Context Test, and the Proverbs Test all loaded on this factor whereas the Achievement score from the Tower Test, the Word Context Test and the Proverbs Test loaded only slightly for the 20- to 49-year-olds and not really at all for the youngest group, the 8- to 19-year-olds. The two tests that anchor this factor, the Color–Word Test and the Trail Making Test, measure the ability to inhibit an overlearned response and flexibility of thinking on a visual–motor task, respectively. This factor was therefore labeled Inhibition.

As is evident in Tables 2 and 3, some of the individual loadings of the various D-KEFS achievement scores on the three factors for the two younger age-groups were quite low. For example, Twenty Questions, Tower Test: Accuracy Ratio, and the Word Context Test all loaded below .30 on all of the three factors for these two younger groups. For the older group, however, only Twenty Questions evidenced weak loadings (see Table 4). This may indicate that not all of the D-KEFS achievement scores are reflective of higher-order executive functioning, although this appears to be less true for the oldest group (50- to 89-year-olds).

Interfactor correlations are presented in Table 5. Moderate interfactor correlations were found across ages, with correlation estimates appearing to increase somewhat with

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Table 6. Study 1: Fit Indices for Measurement Invariance Models

Model	χ^2	ln(*L*)	*k*	RMSR	BIC	DIC
Weak invariance models						
Full loading variant, factor correlations variant	900.597	–450.299	132	0.027	927.897	806.597
Full loading invariant, factor correlations variant	1383.767	–691.884	51	0.061	876.410	829.544
Strict invariance models						
Full loading variant, factor correlations invariant	913.980	–456.990	126	0.028	912.879	797.093
Full loading invariant, factor correlations invariant	1422.613	–711.307	45	0.063	874.124	832.772
Partial invariance models						
Partial loading variant, factor correlations variant	994.863	–497.432	102	0.028	866.485	772.753
Partial loading variant, factor correlations invariant	1022.081	–511.041	96	0.030	858.385	770.167

Note: ln(*L*) denotes log-likelihood; *k* = the number of parameters; RMSR = root mean square residual; BIC = Bayesian information criterion; DIC = Draper's information criterion. *N* = 1,389.

Table 7. Study 1: Best-Fitting Multigroup Model Parameter Estimates

	Conceptual Flexibility			Monitoring			Inhibition		
D-KEFS Achievement Test	8-19	20-49	50-89	8-19	20-49	50-89	8-19	20-49	50-89
Sorting Cond. 1: Free Sort	***.95***	***.95***	***.95***	*.00*	*.00*	*.00*	*.00*	*.00*	*.00*
Sorting Cond. 2: Free Sort Description	***.96***	***.96***	***.96***	*–.00*	*.00*	*.00*	*.01*	*.01*	*.01*
Sorting Cond. 3: Sort Recognition	***.58***	***.58***	***.58***	*.13*	*.13*	*.13*	*.17*	*.17*	*.17*
Trail Making Test	.09	.19	.27	.10	.17	.19	.40	.44	.46
Color–Word Test: Inhibition	*–.03*	*–.03*	*–.03*	*.08*	*.08*	*.08*	***.72***	***.72***	***.72***
Color–Word Test: Inhibition/Switching	*.00*	*.00*	*.00*	*.00*	*.00*	*.00*	***.70***	***.70***	***.70***
Twenty Questions Test	.23	.18	.28	.02	.18	.15	.21	.15	.04
Verbal Fluency: Letter Fluency	.10	.12	.26	.43	.39	.24	.23	.32	.37
Verbal Fluency: Cat. Fluency	.07	.09	.17	.51	.54	.32	.20	.27	.38
Verbal Fluency: Cat. Switch Total	*.00*	*.00*	*.00*	***.90***	***.90***	***.90***	*.00*	*.00*	*.00*
Verbal Fluency: Cat. Switch Accuracy	*–.02*	*–.02*	*–.02*	***.76***	***.76***	***.76***	*.00*	*.00*	*.00*
Design Fluency	.07	.24	.21	.12	.08	.13	.26	.34	.34
Tower Test: Achievement	–.08	.19	.20	.01	–.02	.14	–.08	.22	.43
Tower Test: Accuracy Ratio	.26	–.06	–.19	.26	–.16	.02	.16	.04	–.28
Word Context Test	.24	.27	.46	.18	.27	.14	.18	.22	.26
Proverbs Test	—	.24	.42	—	.31	.17	—	.22	.23

Note: Total *N* = 1,389. Partial loading variant, factor correlations invariant model best fit the data. Parameter estimates reflect factor loadings for each age group: 8- to 19-year olds, 20- to 49-year-olds, and 50- to 89-year-olds. Loading fixed across groups are given in italics. Among those fixed equal across groups, those >.50 are given in boldface italics.

age. For example, the correlation between Cognitive Flexibility and Inhibition in the youngest group was .32 whereas the same correlation in the oldest group was .58.

Invariance models. As stated above, invariance model comparisons were performed in three steps, roughly corresponding to evaluation of weak invariance, strict invariance, and partial invariance models. For partial invariance models, loadings of scores that consistently anchored each factor were constrained to be equal while all other loadings were freed across age groups. Specifically, the loadings constrained to be equal were Sorting Condition 1: Free Sort, Sorting Condition 1: Sort Free Sort Description, Sorting Condition 2: Sort Recognition, Verbal Fluency: Category Switching Total, Verbal Fluency: Category Switching Accuracy, Color–Word Test: Inhibition, and Color–Word Test: Inhibition/Switching.

Fit statistics for the measurement invariance models are presented in Table 6. According to the model fit statistics, the optimal model was one in which factor loadings were partially invariant and factor intercorrelations did not change across age groups (DIC =770.167; BIC = 858.385; root mean square residual [RMSR] = 0.030), corresponding to a form of partial factorial invariance.

Estimates for the best-fitting model are shown in Table 7. In the best-fitting model, the three scores from the Sorting Test were fixed to be equal across groups, anchoring the Conceptual Flexibility factor without cross-loadings. Word Context Test as well as Proverbs Test also evidenced

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

significant loadings (>.40) on this factor in the 50- to 89-year-old group. The second factor, Monitoring, was anchored by the Category Switching Total and Category Switching Accuracy scores from the Verbal Fluency Test whose loadings were fixed across groups. Letter Fluency and Category Fluency scores from the Verbal Fluency Test significantly loaded (>.30) on this factor in all groups as did the Proverbs Test in the 20- to 49-year-old group. The third factor, Inhibition, was again anchored by Color–Word Tests Inhibition and Inhibition/Switching scores whose loadings were constrained to be equivalent across groups. The Trail Making Test also significantly loaded (>.40) on this factor across groups. Inhibition also appeared to be reflected in Verbal Fluency: Letter Fluency and Design Fluency in the 20- to 49-year-olds and in Letter Fluency and Category Fluency from the Verbal Fluency Test, Design Fluency, and the Achievement score from the Tower Test in the oldest group, the 50- to 89-year-olds. Interfactor correlations were constrained to be equal across groups. Conceptual Flexibility was correlated .25 with Monitoring and .30 with Inhibition whereas Monitoring and Inhibition were correlated .21 with one another. Interfactor correlations for this model are included in Table 5.

Study 1 Discussion

Results of exploratory factor analyses of each age group reported in the D-KEFS standardization manual found that a three-factor solution best fit the data for all age groups. A three-factor solution is consistent with previous work by Miyake et al. (2000) using a sample of undergraduates, as well as Lehto et al. (2003) using a sample of children aged 8 to 13 years.

For all groups, the first factor, termed Conceptual Flexibility, was reflected in the three scores from the Sorting Test. This factor appears to represent abilities related to abstraction and concept formation as the Sorting Test was designed to isolate and measure multiple components of concept-formation and problem-solving abilities. In addition to the Sorting Test, the Word Context Test also consistently, albeit weakly, loaded on this factor for all age groups as did the Proverbs Test for the two older age groups.

The second factor, termed Monitoring, comprises prominent loadings on the Category Switching tasks of the Verbal Fluency test. Various studies have demonstrated the importance of switching activities in verbal fluency tasks, and how they are distinct from production activities of the tasks (Bousfield & Sedgewick, 1944; Gruenewald & Lockhead, 1980; Hirshorn & Thompson-Schill, 2006; Nutter-Upham et al., 2008; Rende, Ramsberger, & Miyake, 2002; Smith & Jonides, 1999; Troyer et al., 1997; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). Factor analyses and other studies, for example, suggest that switching and clustering activities during verbal fluency tasks are distinct, and involve different neuropsychological processes (Bousfield & Sedgewick, 1944; Gruenewald & Lockhead, 1980; Nutter-Upham et al., 2008; Troyer et al., 1997). Switching tasks, in particular, are specifically associated with the activity in the lateral prefrontal cortex (Hirshorn & Thompson-Schill, 2006; Smith & Jonides, 1999), as opposed to pure production fluency tasks, which are associated more with temporal lobe activity (Troyer et al., 1998). Consistent with this, performance on D-KEFS Category Switching is more highly correlated with WAIS-IV Working Memory than is Category Fluency, which is more highly correlated with WAIS-IV Verbal Comprehension (Wechsler, 2008). Verbal fluency switching tasks appear more effortful than clustering or production tasks, requiring the monitoring of the current category use. This factor appears to represent abilities closely linked to working memory, as the core of this factor requires active manipulation and simultaneous maintenance of relevant information. Whereas the D-KEFS does not include tests specifically intended to measure working memory, the Category Switching subtests of the Verbal Fluency Test appear to tap working memory abilities, a possibility supported by studies demonstrating that switching in particular in verbal fluency tasks creates prominent working memory loads (Rende et al., 2002; Troyer et al., 1997).

The third factor, termed Inhibition, concerns the ability to deliberately inhibit a dominant or automatic response. This factor is anchored by the Trail Making Test and Color–Word Inhibition Test, a modified version of the procedure originally developed by Stroop (1935) which has been termed the prototypical Inhibition test (Miyake et al., 2000). The Trail Making Test and the Color–Word Test represent core Inhibition tasks as they require the examinee to inhibit one task, connecting numbers in order or reading, and do another, switching to letters or naming the color in which the words are printed, placing demands on inhibitory abilities (Lezak et al, 2004).

It is interesting to compare the factor loading patterns of the Trail Making Test and Color–Word Test to that of the Category Switching test. Although all three tests involve switching tasks, the former tests load on the Inhibition factor, and the latter loads on the Monitoring factor. One difference between the first two tests and the Category Switching test is that the first two involve switching on features of an external stimulus, whereas the latter test involves switching on internal representations of a previously generated response. The stimulus features of the Trail Making Test and the Color–Word Test may act as cues, which in turn elicit automatic, prepotent responses that must be inhibited to respond correctly. In the case of the Category Switching test, in contrast, there are no external stimuli to guide responding, so switching must be done on the basis of information about previous responses, presumably maintained in working memory.

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Table 8. Study 2: Observed Eigenvalues and Monte Carlo *p* Values

	Eigenvalue no.							
	1	2	3	4	5	6	7	8
Observed eigenvalue	4.619	1.920	1.632	1.234	1.010	.853	.737	.709
Monte Carlo *p*	.000	.000	.000	.515	1.000	1.000	1.000	1.000

Note: *p* values were calculated by Monte Carlo methods as described in the text.

Age-associated invariance modeling indicated that intercorrelations between factors do not change significantly with age. Nonetheless, some loadings do change with age, whereas others—namely those that consistently anchor each factor—do not. Examination of interfactor correlations from EFA analyses revealed that correlations between the Inhibition factor and the other two factors were somewhat larger than correlations involving the other factors, suggesting that Inhibition may more strongly reflect a general superordinate EF factor than the other factors. Although not significant in model comparisons, this appeared to be especially evident in the older groups, which is broadly consistent with previous research by Hull et al. (2008) in suggesting that EF abilities are less differentiated in older individuals.

Study 2

To generalize the finding of a three-factor solution for the D-KEFS, we replicated the exploratory structural analyses in a second, independent sample. Data for Study 2 consisted of data collected as part of the Iowa Youth Development Project (I-YDP), a larger investigation of relationships between adolescents, their parents, and adolescent psychopathology. As adolescent males are more likely to exhibit greater variance in internalizing and externalizing behaviors than females, data collection resources were focused on males. Participants included 174 male adolescents aged 11 to 16 years (mean = 13.6, SD = 1.4). One adolescent was color blind and was therefore not able to be administered the Color–Word Test. The I-YDP used a broad-based sampling strategy to accrue a sample of normally developing urban and rural adolescents representative of the Midwestern area in which participants were recruited. Participants were recruited from a child participant pool maintained at the University of Iowa in addition to fliers in the community or via an advertisement placed in the daily newsletter of the local university hospital. To assure a typically developing sample, exclusion criteria included the following: mental retardation, autism spectrum disorder, neurological disorder, past head injury requiring hospitalization, life-threatening medical illness, having been held back a grade, and being diagnosed with a reading disorder. The eight D-KEFS tests suitable for the age-range of the sample were administered during a single session.

Analyses

Analyses were similar to those performed in Study 1. Exploratory factor models were fit to the raw data using maximum likelihood estimation. Models comprising between one and eight factors were fit, with loading matrices rotated using quartimin rotation. Monte Carlo *p* values (i.e., parallel analysis; Horn, 1965) and Velicer's MAP criterion (Velicer et al., 2000) were used to determine the optimal number of factors to retain.

Results

Consistent with the results of Study 1, parallel analyses suggested a three-factor solution best fit the data. Monte Carlo *p* values and observed eigenvalues are presented in Table 8. MAP analyses suggested a one-factor solution. Given the correspondence of these results with those in Study 1, it was determined that a correlated three-factor solution provided optimal fit to the data for each of the three groups. Results of the three-factor EFA are presented in Table 9.

Consistent with analyses done in the standardization sample, the first factor, Conceptual Flexibility, was anchored by the three scores from the Sorting Test: Free Sort, Free Sort Description, and Sort Recognition. Additionally, the Word Context Test loaded on this first factor. The second factor, Monitoring, was again anchored by the two Category Switching scores from the Verbal Fluency Test. Additionally, the other two Verbal Fluency scores, Letter Fluency and Category Fluency, as well as the Twenty Questions Test loaded on this factor. Lastly, the third factor, Inhibition, was again anchored by the Inhibition and Inhibition/Switching scores from the Color–Word Test in addition to the Trail Making Test. Additionally, Letter Fluency, and both Tower Test scores loaded most highly on this factor. Again, moderate inter-factor correlations were found and similar to the previous study, the largest correlation was found between Inhibition and Monitoring. Conceptual Flexibility was correlated .28 with Monitoring and .31 with Inhibition and Monitoring was correlated .35 with Inhibition.

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Table 9. Study 2: Quartimin Rotated Exploratory Factor Analysis for I-YDP Data

D-KEFS Achievement Tests	Conceptual Flexibility	Monitoring	Inhibition	h^2
Sorting Cond. 1: Free Sort	**.95**	–.03	.04	.14
Sorting Cond. 2: Free Sort Description	**.91**	.02	–.00	.16
Sorting Cond. 3: Sort Recognition	**.53**	.08	–.20	.57
Trail Making Test	.27	.04	**–.44**	.65
Color–Word Test: Inhibition	–.00	.02	**–.79**	.37
Color–Word Test: Inhibition/Switching	–.04	–.03	**–.79**	.41
Twenty Questions Test	.09	**.33**	.01	.87
Verbal Fluency: Letter Fluency	.08	.21	**–.32**	.77
Verbal Fluency: Cat. Fluency	–.05	**.37**	–.20	.79
Verbal Fluency: Cat. Switch Total	.01	**.97**	–.02	.04
Verbal Fluency: Cat. Switch Accuracy	–.00	**.98**	.03	.07
Design Fluency	–.01	.21	**–.39**	.76
Tower Test: Achievement	.17	–.15	**–.34**	.85
Tower Test: Accuracy Ratio	.12	–.07	–.27	.91
Word Context Test	**.37**	.23	–.05	.75

Note: I-YDP = Iowa Youth Development Project. N = 174 except for color–word test where N = 173. h^2 = unique variance. Loadings ≥.30 are given in boldface.

Study 2 Discussion

Results obtained from data from the first study, using the D-KEFS standardization sample, were replicated in this independent community sample of adolescent males. That is, a similar three factor solution obtained from the standardization sample was found to best fit these independent data, providing convergent validity that the Total Achievement Scores of the D-KEFS can be accounted for by three interrelated factors.

General Discussion

The current investigation represents the first examination of the factor structure and age-related factorial invariance of the D-KEFS. Our results suggest that a three-factor solution is optimal for the D-KEFS. This solution was found across the three age groups reported in the standardization manual as well as in an independent community sample of male early adolescents. The three factors found across groups and samples were labeled Conceptual Flexibility, Monitoring, and Inhibition.

When the age-related factorial invariance of these factors was examined, it was found that the best fitting unrestricted confirmatory model fixed some factor loadings while freeing others, and constrained factor intercorrelations to be consistent across groups. This indicates that, whereas there is age-related invariance in the scores that anchor each factor, there is age-related variance in the factor loading pattern for other achievement scores. This is consistent with results of our EFA analyses in which anchor scores consistently and highly loaded on the same factor across groups whereas other scores loaded slightly if at all on factors and did not often consistently load on the same factor across groups.

Contributions to the Understanding of D-KEFS Measurement Properties

Three-factor structure. The present study greatly contributes to our understanding of the D-KEFS. Results of the present study suggest that D-KEFS achievement scores can be conceptualized in a three-factor space with the Sorting Test, Verbal Fluency Category Switching Test, and the Color–Word Inhibition Test anchoring each of the factors, respectively. This same structure is consistent across the age groups of the standardization sample and also emerges when data from a separate, independent sample are examined. Although this convergent validity in an independent, community sample is of great importance, it should be noted that this independent sample consisted of typically developing, early adolescent males. One possible explanation for the slightly divergent loadings found in this independent sample related to its gender homogeneity. Whereas the D-KEFS standardization sample were of mixed gender examinees, this second, independent sample consisted of entirely males. It is therefore important for future research to replicate these findings in more diverse samples of both males and females as well as samples that include a wider age-range of participants.

Identifying a replicable three-factor structure is important, as it helps clarify the factor structure of the D-KEFS. For example, speculation by the authors that the D-KEFS might primarily reflect verbal and nonverbal processing forms of EF (Delis et al., 2001a) were not supported by the data. Conversely, however, the three-factor structure of D-KEFS does parallel previous investigations of EF structure,

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

as do our findings that these three factors broadly cohere as a set of correlated abilities. Recent work (Latzman, Elkovitch, Young, & Clark, in press) provides evidence for the validity of these three factors as they have been shown to predict relevant outcomes in distinct ways.

Age-related measurement properties. Results of the current study also provide insight into how measurement properties of the D-KEFS are associated with age. Although the basic three-factor structure appears to be invariant across age groups, and major indicators of these factors appear to be stable across ages, current results highlight that secondary measures of the factors change in their measurement properties across age groups. Model comparisons also indicate that relationships between the factors are largely stable across age, and that the Inhibition factor in particular may more strongly reflect higher order EF functioning. Consistent with the work of Hull et al. (2008), who found that EF becomes less differentiated in older individuals, there was some evidence from age-specific analyses that factors may be more strongly related in the oldest age groups, although this trend was not significant in overall model comparisons. Further research is needed to understand the D-KEFS and EF more generally in old age.

Parallels With General Models of EF Structure

The structure of the D-KEFS as identified here parallels findings by Miyake et al. (2000) in an adult sample as well as by Lehto et al. (2003) in a sample of children in suggesting a three-factor structure among measures of EF. It should be noted that whereas the D-KEFS consists of some tests similar to those used by Lehto et al. and Fisk and Sharp (2004), other than the Stroop, the tests included in the D-KEFS are entirely different from those used by Miyake et al. (2000). Nonetheless, a highly similar three-factor solution emerged from the data in the present study as the structure previously found by these researchers. Additionally, consistent with previous work, correlations between the factors suggest that a higher order EF domain can be identified.

In a study of normally developing college students based on lower level cognitive tasks, Miyake et al. (2000) confirmed three distinct yet related EF factors: Shifting, Updating, and Inhibition. The Shifting factor described by Miyake et al. (2000), reflected in Wisconsin Card Sorting (WCST) performance, requires shifting back and forth between multiple tasks or operations and performing new operations while being faced with proactive interference. This factor strongly corresponds to the Conceptual Flexibility identified in our data, reflected in the Sorting Test requirements of shifting mental sets, initiating problem solving behavior, formulating new concepts, inhibiting previous sorts and description responses, and generally engaging in flexible behavior and thinking.

The Updating factor described by Miyake et al. (2000) requires monitoring and evaluating the relevance of new information for the task at hand and, if appropriate, updating information in working memory with the new information if appropriate. We decided to term the closely related factor that emerged in our data Monitoring to better reflect the specific importance of actively monitoring and evaluating information in working memory that the D-KEFS Category Switching scores reflect. Although these tasks were not originally intended to measure working memory, per se, these abilities are closely linked to working memory abilities. Switching in verbal fluency tasks creates prominent working memory loads (Rende et al., 2002; Troyer et al., 1997) and, as such, has been shown to be associated with the lateral portion of the prefrontal cortex (Hirshorn & Thompson-Schill, 2006; Smith & Jonides, 1999), consistent with Miyake et al.'s Updating factor.

Miyake et al.'s (2000) third factor, Inhibition, reflects the ability to deliberately control prepotent responses, responses that are automatic and dominant. These abilities are most clearly expressed in performance on the Stroop task (Miyake et al., 2000). This is clearly reflected in our data in the third factor, which we, not surprisingly, also termed Inhibition. In our data, this factor is anchored by scores from both the Inhibition as well as the Inhibition/Switching tasks on the Color–Word Test, a version of the Stroop task, which is the only EF task included in the D-KEFS that overlaps with those used by Miyake et al.

Results of the current investigation are also consistent with Miyake et al.'s (2000) assertion that there are both unitary and nonunitary components of the executive system. This is reflected in findings that EFs are separable but moderately correlated. Data from diverse samples (Fisk & Sharp, 2004; Hull et al., 2008; Lehto et al., 2003), as well as data from the two independent samples reported in the current study, are consistent with this assertion. In our data, we find moderate interfactor correlations, reflecting a unified, higher order structure that is, nonetheless, separable at a lower order level. This organization parallels findings that the frontal region comprises a network of dissociable but linked brain regions (Lezak et al., 2004), consistent with the aims of the D-KEFS to measure frontally mediated functions (Delis et al., 2001a).

Limitations

Although providing an important contribution to the literature concerning the D-KEFS, our age-related analyses were limited to the broad age groups as well as to achievement scores for which data was provided in the standardization manual. Examination of more detailed age-related differences was thus not possible. Additionally, the present study

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

examined data from the mixed-gender standardization sample as well as an independent sample of typically developing adolescent males. It is therefore not known how gender may affect our findings. Furthermore, a sample of neurologically impaired participants or participants with different forms of psychopathology was not examined. Future studies are needed to examine the structure of the D-KEFS across genders as well as in various identified groups.

Conclusions

Despite these limitations, our results help begin to clarify the measurement properties of a widely used neuropsychological inventory, the D-KEFS, in particular, and also contribute to the ongoing discourse concerning the nature of EF, generally. Our data indicate that performance on the D-KEFS reflects three prominent factors—Conceptual Flexibility, Monitoring, and Inhibition—that represent dissociable but related aspects of general EF functioning. These results are broadly consistent with previous investigations of EF structure, and will help clarify the underlying mechanisms of EF in future work.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research and/or authorship of this article:

The Iowa-Youth Development Project was supported, in part, by the University of Minnesota Press, the American Psychology-Law Society, and the American Academy of Forensic Psychology, awarded to R. D. Latzman.

Notes

1. More simple-structure models, such as freeing loadings that appeared significant in the EFAs were tried, but did not provide adequate fit.
2. In analyses in which different combinations of variables were included, including removing highly correlated measures, the only exception to the three-factor solution suggested by parallel-analysis was in the oldest group (50-89 year olds) where parallel analysis suggested a two-factor solution, consistent with the results of Hull et al. (2008). For the sake of completeness, and because results were similar either way, we report results on the complete set of measures.

References

Benton, A. L., & Hamsher, K. deS. (1976). *Multilingual aphasia examination manual.* Iowa City: University of Iowa.

Baron, I. S. (2003). *Neuropsychological evaluation of the child.* New York: Oxford University Press.

Baron, I. S. (2004). Delis-Kaplan Executive Function System. *Child Neuropsychology, 10*, 147-152.

Brown, R. R., & Partington, J. E. (1942). The intelligence of the narcotic drug addict. *Journal of General Psychology, 26*, 175-179.

Bousfield, W. A., & Sedgewick, H. W. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology, 30*, 149-165.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive functions* (pp. 81-116). Hove, UK: Psychology Press.

Delis, D. C., Kaplan, E., & Kramer, J. (2001a). *Delis-Kaplan Executive Function System.* San Antonio, TX: Psychological Corporation.

Delis, D. C., Kaplan, E., & Kramer, J., (2001b). *Delis-Kaplan Executive Function System (D-KEFS) technical manual.* San Antonio, TX: Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Holdnack, J. (2004). Reliability and validity of the Delis-Kaplan Executive Function System: An update. *Journal of the International Neuropsychological Society, 10*, 301-303.

Delis, D. C., Squire, L. R., Bihrle, A., & Massman, P. (1992). Component analysis of problem-solving ability: Performance of patients with frontal lobe damage and amnesic patients on a new sorting test. *Neuropsychologia, 13*, 135-143.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30*, 257-303.

Duncan, J., Johnson, R., Swales, M., & Freer, C. (1997). Frontal lobe deficits after head injury: Unity and diversity of function. *Cognitive Neuropsychology, 14*, 713-741.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). New York: Cambridge University Press.

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26*, 874-890.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*, 31-60.

Greve, K. W., Farrell, J. F., Besson, P. S., & Crouch, J. A. (1995). A psychometric analysis of the California Card Sorting Test. *Archives of Clinical Neuropsychology, 10*, 265-278.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 225-240.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7*, 191-205.

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010

Hirshorn, E. A., & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia, 44*, 2547-2557.

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

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*, 2017-2036.

Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology, 22*, 508-522.

Jöreskog, K. G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models.* Cambridge, MA: Abt Books.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review, 17*, 213-233.

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (in press). The contribution of executive functioning to academic achievement among male adolescents. *Journal of Clinical and Experimental Neuropsychology.*

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology, 49A*, 29-50.

Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59-80.

Levin, H. S., Fletcher, J. M., Kufera, J. A., Harward, H., Lilly, M. A., Mendelsohn, D., et al. (1996). Dimensions of cognition measured by the Tower of London and other cognitive tasks in head-injured children and adolescents. *Developmental Neuropsychology, 12*, 17-34.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Lowe, C., & Rabbitt, P. (1997). Cognitive models of aging and frontal lobe deficits. In P. Rabbitt (Ed.), *Methodology of frontal and executive functions* (pp. 39-59). Hove, UK: Psychology Press.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and disunity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Muthén, L. K., & Muthén, B. O. (1998-2007). *Mplus user's guide* (5th ed.). Los Angeles: Muthén & Muthén

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.

Newcombe, F. (1969). *Missile wounds of the brain.* New York: Oxford University Press.

Nutter-Upham, K. E., Saykin, A. J., Rabin, L. A., Roth, R. M., Wishart, H. A., Pare, N., et al. (2008). Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. *Archives of Clinical Neuropsychology, 23*, 229-241.

Reitan, R. M., & Wolfson, D. (1993). *Halstead-Reitan Neuropsychological Battery.* Tuscon, AZ: Neuropsychology Press.

Rende, B., Ramsberger, G., & Miyake, A. (2002). Commonalities and differences in the working memory components underlying letter and category fluency tasks: A dual-task investigation. *Neuropsychology, 16*, 309-321.

Satorra, A. (2002). Asymptotic robustness in multiple group linear-latent variable models. *Econometric Theory, 18*, 297-312.

Savalei, V. (2008). Is the ML chi-square ever robust to nonnormality? A cautionary note with missing data. *Structural Equation Modeling, 15*, 1-22.

Skunk, A. W., Davis, A. S., & Dean, R. S. (2006). Review of Delis-Kaplan Executive Function System (D-KEFS). *Applied Neuropsychology, 13*, 275-279.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*, 1657-1661.

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology, 18*, 643-662.

Swanson, J. (2005). The Delis-Kaplan Executive Function System: A review. *Canadian Journal of School Psychology, 20*, 117-128.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499-504.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35*, 427-449.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas Jackson at seventy* (pp. 41-71). Boston: Kluwer.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV) technical and interpretative manual.* San Antonio, TX: NCS Pearson.

Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive functions across the life span. *Acta Psychologica, 115*, 167-183.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.

Downloaded from http://asm.sagepub.com at UNIV NEBRASKA LIBRARIES on July 13, 2010