



Published in final edited form as:

J Exp Psychol Gen. 2019 December ; 148(12): 2104–2119. doi:10.1037/xge0000589.

Integrating Verbal Fluency with Executive Functions: Evidence from Twin Studies in Adolescence and Middle Age

Daniel E. Gustavson,

Department of Psychiatry, Center for Behavior Genetics of Aging, University of California, San Diego

Matthew S. Panizzon,

Department of Psychiatry, Center for Behavior Genetics of Aging, University of California

Carol E. Franz,

Department of Psychiatry, Center for Behavior Genetics of Aging, University of California, San Diego

Chandra A. Reynolds,

Department of Psychology, University of California, Riverside

Robin P. Corley,

Institute for Behavioral Genetics, University of Colorado Boulder

John K. Hewitt,

Institute for Behavioral Genetics and Department of Psychology and Neuroscience, University of Colorado Boulder

Michael J. Lyons,

Department of Psychological and Brain Sciences, Boston University

William S. Kremen*,

Department of Psychiatry, Center for Behavior Genetics of Aging, University of California, San Diego and Veterans Affairs San Diego Healthcare System

Naomi P. Friedman*

Institute for Behavioral Genetics and Department of Psychology and Neuroscience, University of Colorado Boulder.

Abstract

The relationship of verbal fluency to executive functions (EFs) remains somewhat unclear. Verbal fluency is sometimes considered an EF ability, but is not often included in the same models as other well-studied EFs (inhibition, shifting, and working memory updating). We examined the associations between verbal fluency and EFs at two ages with the unity/diversity model, which includes common and domain-specific EF factors. Participants were 813 adolescent twins from the Colorado Longitudinal Twin Sample (mean age 17 years) and 1290 middle-aged twins from the

Correspondence concerning this article should be addressed to Daniel Gustavson, Department of Psychiatry, University of California San Diego, 9500 Gilman Dr. (MC 0738), La Jolla, CA 92093. dgustavson@ucsd.edu.

*William S. Kremen and Naomi P. Friedman should be considered co-senior authors.

Vietnam Era Twin Study of Aging (mean age 56 years) who completed multiple measures of EFs, verbal fluency, vocabulary, and nonverbal cognitive ability. Results revealed that, in both samples, a General Fluency factor (i.e., comprising both phonemic and semantic fluency measures) was associated with the Common EF factor, but also with variance unique to working memory updating, working memory span, and set-shifting. In adolescents, semantic fluency also had unique associations with shifting beyond its shared variance with phonemic fluency and Common EF. After accounting for EFs and other cognitive abilities, there were unique genetic and environmental influences on the General Fluency and Semantic-Specific latent factors. These results suggest that verbal fluency ability may best be viewed as an amalgamation of general EF variance (i.e., Common EF ability), variance shared with other EFs (e.g., Updating-Specific ability), and multiple sources of unique genetic/environmental variance (i.e., General Fluency and Semantic-Specific abilities). These associations between verbal fluency and EFs generalize to populations that differ in age by approximately 40 years.

Keywords

heritability; word fluency; category fluency; executive control; twin study

Introduction

Verbal fluency is often described as an executive function (EF) alongside prepotent response inhibition (inhibition), task-set shifting (shifting), and working memory (WM) updating (Delis, Kaplan, & Kramer, 2001; Diamond, 2013; Henry & Crawford, 2004; Phillips, Bull, Adams, & Fraser, 2002). However, its relations to other well-studied EFs has received little attention (Shao, Janse, Visser, & Meyer, 2014; Unsworth, Spillers, & Brewer, 2011; Whiteside et al., 2016), especially regarding the general EF processes accounting for performance across multiple situations (Miyake & Friedman, 2012). The goal of the current study is to quantify how individual differences in verbal fluency are associated with other EFs using two large genetically informative twin samples. Specifically, we ask whether fluency might be considered a separable EF or a composite of other well-studied EFs. We answer this question in part by examining evidence for distinct genetic influences on fluency beyond those for EFs.

Unity and Diversity of Executive Functions

In general, correlations among EF tasks are low to moderate (e.g., Miyake et al., 2000). EF processes must be assessed within a specific task context (e.g., naming the ink of a color word in the Stroop task) which necessarily includes variance attributable to non-EF processes (e.g., color processing, articulation speed), even when adequately controlling for baseline conditions (Miyake & Friedman, 2012). This problem of *task impurity* has made it difficult for researchers to elucidate the commonality and separability of various EF processes. However, latent variables models have advanced our understanding of EFs substantially by allowing researchers to isolate the general variance underlying performance across all three EF domains (called “Common EF”) from domain-specific EF variance (e.g.,

“Updating-Specific” or “Shifting-Specific” abilities) and task-specific variance (Friedman & Miyake, 2017; Friedman et al., 2016; Friedman et al., 2008; Miyake & Friedman, 2012).

Common EF represents the ability to activate, maintain, and implement goal-directed thoughts and actions, as these general goal-management processes are needed across multiple situations (Friedman & Miyake, 2017). In contrast, Updating-Specific ability is thought to reflect the gating of information in WM, controlled by the basal ganglia (e.g., Frank, Loughry, & O’Reilly, 2001), and may also involve retrieval from episodic memory. The Shifting-Specific factor is thought to reflect the ability to flexibly replace goals that are no longer necessary (Friedman & Miyake, 2017; Miyake & Friedman, 2012). Moreover, Shifting-Specific ability has been found to be negatively associated with concurrent intelligence (Friedman et al., 2008) and earlier self-restraint (Friedman, Miyake, Robinson, & Hewitt, 2011), suggesting a potential stability-flexibility tradeoff (Friedman & Miyake, 2017; Goschke, 2000; Herd et al., 2014). As yet, there is no evidence for Inhibition-Specific variance, suggesting that individual differences in response inhibition are entirely captured by Common EF (Friedman et al., 2016; Gustavson, Panizzon, Franz, et al., 2018; Ito et al., 2015), and that not all EFs have unique variance components.

Latent variable studies using this unity/diversity framework have also utilized twin samples. Twin studies can be used to identify the extent to which individual differences (i.e., variance) and correlations among traits (i.e., covariance) are explained by genetic versus environmental influences. For example, prior studies suggest that there are both common genetic influences that unite different EFs, as well as genetic influences that are unique to particular EFs (Engelhardt, Briley, Mann, Harden, & Tucker-Drob, 2015; Friedman et al., 2016; Friedman et al., 2008; Gustavson, Panizzon, Franz, et al., 2018). The Common EF factor shows strong phenotypic and genetic stability (phenotypic correlations $r_s=.86-.97$, genetic correlations, or $r_g=1.0$) from adolescence to early adulthood (Friedman et al., 2016) and across middle age (Gustavson, Panizzon, Elman, Franz, et al., 2018). Strong phenotypic and genetic stability were observed for the domain-specific EF factors in both studies. With respect to verbal fluency, a pertinent question is whether fluency is distinguished from other EFs by unique genetic and/or environmental influences.

Overlap Between EFs and Verbal Fluency

Verbal fluency — the ability to rapidly generate exemplars from a cued category such as “types of animals” — is often included as an EF in reviews and meta-analyses (e.g., Alvarez & Emory, 2006; Snyder, 2013). However, it remains unclear how it fits in with other well-studied EFs, as research on EF has been dominated by inhibition, shifting, and WM updating (Miyake & Friedman, 2012). The inclusion of verbal fluency in the domain of EF may stem partly from its obvious face validity: Controlled retrieval of words involves establishing a task set, strategic search, and avoidance of repetition (Fisk & Sharp, 2004; Shao et al., 2014), processes that appear similar to those involved in other EFs (Diamond, 2013; Miyake & Friedman, 2012). Its inclusion as an EF ability may also be related to historical studies of fluency impairments in frontal-lobe patients (Henry & Crawford, 2004), similar to other neuropsychological measures of EFs (Alvarez & Emory, 2006), suggesting the two draw on at least some similar neural substrates.

Some empirical work supports verbal fluency's overlap with tasks from other well-studied EFs. One recent study that found that a fluency latent variable correlated with inhibition, shifting, and WM updating latent variables (Hedden & Yoon, 2006; see also Shao et al., 2014; Whiteside et al., 2016). A similar study showed that a latent factor comprising phonemic and semantic tasks was positively associated with inhibition, WM capacity, vocabulary, and processing speed, although only WM capacity and vocabulary remained significant in the multiple regression model (Unsworth et al., 2011).

This work provides some validation of the overlap between fluency and EFs. However, given the substantial correlations among EF factors, it remains unclear whether the associations between verbal fluency and other EFs are explained by Common EF variance, or if performance on verbal fluency tasks also rely on Updating-Specific or Shifting-Specific abilities. In fluency tasks, Common EF ability is likely involved in maintaining the goal representation, adhering to rules, using the prompt to guide lexical search, and avoiding interference from associated words. However, some of these processes may rely on Updating-Specific ability, such as employing controlled retrieval from long-term memory (Miyake & Friedman, 2012). Similarly, because performing well on fluency tasks may involve strategic switching between sub-categories (Troyer & Moscovitch, 2006), it is possible that Shifting-Specific ability contributes to fluency performance as well. To the extent that Updating- or Shifting-Specific abilities are utilized in fluency tasks, then fluency is not comparable to WM updating or shifting (e.g., relying on only Common EF and Fluency-Specific abilities), but would rather reflect a combination of other general and specific EF processes.

Another factor that complicates understanding of fluency's position as an EF is that fluency is itself multifaceted. Measures of phonemic fluency (e.g., naming words that begin with *F*) and semantic fluency (e.g., naming types of animals) correlate with one another and with vocabulary (Whiteside et al., 2016), but there is mounting meta-analytic evidence that phonemic and semantic fluency are differentially associated with neurodegenerative and psychiatric conditions such as Alzheimer's disease (Henry & Crawford, 2004, 2005, 2005; Henry, Crawford, & Phillips, 2004), and may rely on somewhat distinct brain regions (Henry & Crawford, 2004).

Our recent work using one of the samples described here (the Vietnam Era Twin Study of Aging; VETSA) supported a two-factor model of fluency similar to that of the unity/diversity model of EFs (Gustavson, Panizzon, Elman, Beck, et al., 2018). In that analysis, variance in phonemic and semantic fluency subtests were best modeled with two orthogonal latent factors: a General Fluency factor accounted for variance in all six subtests, and a Semantic-Specific factor accounted for variance in semantic fluency beyond that variance captured by the general factor. All the covariance among phonemic fluency tasks were captured by the general factor. This model was supported at two time-points (mean age 56 and 62), with considerable stability in individual differences across these ages. Similar to other EFs, there were considerable genetic influences on both factors, suggesting that even within the domain of verbal fluency, there appear to be both common and unique sources of genetic variance.

Integrating verbal fluency with the unity/diversity model of EFs may therefore elucidate the nature of verbal fluency as an EF process, further refine theoretical models of verbal fluency, and better illuminate its differential association with neurodegenerative and psychiatric disorders. Furthermore, there is not necessarily a one-to-one correspondence between the phenotypic and underlying genetic/environmental factor structures of cognitive abilities (Kremen et al., 2009; Vasilopoulos et al., 2012), so examining these questions in the context of a twin study provides additional information about their structure and covariance.

Hypotheses of the Current Study

We expected to observe one of three potential scenarios when examining the association between verbal fluency and EFs. These scenarios are displayed in Figure 1 (panels A-C); Figure 1D displays a simplified version of the Cholesky decomposition used to evaluate these scenarios in the genetic analyses.

In the first scenario, verbal fluency is viewed as a separable EF factor, akin to WM updating or shifting. We expected that much of the variance in the General Fluency factor would be associated with the Common EF factor (path a in Figure 1D), but there would also be unique genetic and/or environmental variance (path c) and no association with domain-specific EF factors (path b). In the second scenario, fluency ability is viewed as a combination of variance in other EF processes. If so, we expected that the General Fluency factor would be associated with the Common EF factor (path a) as well as with domain-specific EF factors (path b) including Updating-Specific or Shifting-Specific factors (in adolescents) and a WM-Specific factor (in middle-aged adults).¹ Furthermore, after controlling for EFs, we expected little to no unique genetic variance in the General Fluency factor (path c). The third scenario was a combination of the first two. This pattern would suggest that verbal fluency requires both general (path a) and specific EF abilities (path b), as well as other processes (path c).

In all scenarios, we expected no correlation between the Semantic-Specific factor and the Common EF factor because the General Fluency factor should already capture common variance among fluency tasks, including variance due to general EF abilities. Moreover, to the extent that unique variance in either fluency factor represents a unique EF process, we expect them to be explained primarily by genetic influences rather than environmental influences, as we have observed for the other domain-specific EF factors at both ages.

The current study examined these associations in two twin samples of adolescents (mean age 17 years) and middle-aged adults (mean age 56). Examining these hypotheses at both ages is useful for replicating the basic results, but also informs the generalizability of these findings. We have shown that the unity/diversity model of EFs is supported at both ages despite the use of rapidly-paced computerized tasks in adolescents and neuropsychological measures in middle-aged adults (Friedman et al., 2008; Gustavson, Panizzon, Franz, et al., 2018). If similar patterns are observed for fluency, including their associations with EFs, it will

¹Although both Updating-Specific (adolescents) and WM-Specific (middle-aged adults) factors refer to unique variance in WM processes not captured by Common EF, these factors are distinguished because the WM tasks in adolescents had more complex WM updating demands than the WM span tasks administered in middle-aged adults. However, WM updating and WM span are quite similar (Schmiedek, Hildebrandt, Lövdén, Wilhelm, & Lindenberger, 2009), even at the genetic level (Friedman & Hewitt, 2013).

provide further evidence that there is a similar structure of EFs at key times in the lifespan when mean-level performance is changing. Examining these associations at the level of latent variables also provides a better estimate of the true heritability and genetic covariance among traits, as measurement error is included in the estimates of nonshared environmental influences on individual tasks.

Method

Sample characteristics for both samples are displayed in Table 1. All studies were approved by the institutional review board at participating institutions.

Sample 1: Adolescents

Participants.—Analyses were based on 813 individuals (420 females, 393 males) from same-sex twin pairs (216 full monozygotic [MZ] twin pairs, 185 full dizygotic [DZ] twin pairs, and 9 unpaired twins) in the Colorado Longitudinal Twin Sample (CLTS). Participants completed the measures of verbal fluency between ages 16 and 17 ($N = 811$, $M = 16.58$ years). Measures of EF were completed approximately one year later ($M = 17.25$ years). Most twins completed both waves of assessment ($N = 786$). The CLTS sample was recruited through the Colorado Department of Health based on twins born between 1984 and 1990 and is representative of the population of Colorado at that time. Participants identified as white (80.9%), Hawaiian or Pacific islander (0.2%), American Indian or Alaskan (1.1%), more than one race (4.7%), or unknown or unreported race (13.2%). Hispanic individuals comprised 9.5% of the sample. For more information on the sample characteristics, see Rhea, Gross, Haberstick, and Corley (2013).

Measures.

Verbal Fluency. —Verbal fluency was measured using two phonemic and two semantic fluency tasks. In the phonemic fluency tasks (*S-P* and *G-T*), participants had three minutes per trial to write as many words as possible beginning with one letter and ending with another (e.g., beginning with *S* and ending with *P*). In the semantic fluency tasks, participants had three minutes per trial to list as many names of “things that are often round” and “things that are often metal.” These measures were performed as a part of a larger “Specific Cognitive Abilities” battery developed for the Colorado Adoption Project (Defries, Plomin, Vandenberg, & Kuse, 1981) based on the battery employed in the earlier Hawaii Family Study of Cognition (DeFries et al., 1974). Semantic fluency tasks were completed together, and phonemic tasks were completed together later in the battery (separated by card rotation and mathematic operations tasks).

Executive function.—Details and scoring procedures for the nine computerized EF tasks have been described in previous work (Friedman et al., 2008) but are summarized here. All accuracy data were arcsine transformed to improve normality. Reaction time (RT) data were analyzed excluding errors (and trials after errors for switching tasks only), and after applying a within-subject trimming procedure (Wilcox & Keselman, 2003).² To reduce the influence of between-subject extreme scores, we replaced task scores farther than three standard deviations from the group mean with values three standard deviations from the mean (<2.1%

of all observations). Scores for the RT measures (i.e., stop signal, Stroop, all switch costs) were multiplied by -1 in all analyses so higher numbers always reflected better performance.

The three inhibition tasks required participants to stop a dominant or prepotent response. In the antisaccade task (Roberts, Hager, & Heron, 1994), participants saw cues (black squares) quickly flash on the left or right side of the screen. They had to avoid the reflexive tendency to saccade to these cues and instead immediately look to the opposite side of the screen to identify a digit that appeared briefly before being masked. The dependent measure was the total accuracy across 90 target trials.³ In the stop-signal task (Logan, 1994), participants saw a green arrow and were instructed to quickly indicate its direction (using the left or right keys), but withhold responses on the 25% of trials where the color of the arrow changed to red. The dependent measure was the stop-signal RT, an estimate of the subjects' stopping process based on how the program adapted to their responses over the course of three blocks of 96 trials each. In the Stroop task (Stroop, 1935), the dependent measure was the RT interference, subtracting RTs across 60 incongruent trials (e.g., *RED*, printed in blue) from RTs on 60 baseline trials with colored strings of asterisks.

The three shifting tasks required participants to switch between categorization dimensions according to a cued stimulus. The dependent measures were local switch costs: the average RT during trials that required a switch between categorization rules minus the average RT during trials where the same rule was repeated (across two blocks of 48 trials each). In the number-letter task (Rogers & Monsell, 1995), subjects judged whether number-letter pairs contained a vowel or consonant or contained an odd or even number, based on where the stimulus appeared in a quadrant of a square on the screen (top or bottom). In the color-shape task (Miyake, Emerson, Padilla, & Ahn, 2004), subjects judged whether colored shapes were either green or red or were a circle or triangle, based on a cue that appeared 150 ms before the word appeared (a *C* or an *S*) and remained on screen until the response. In the category-switch task (Mayr & Kliegl, 2000), subjects judged whether words described something living or non-living or something bigger or smaller than a soccer ball, based on a similar cue (a heart or a cross).

The three WM updating tasks required participants to monitor and continuously manipulate the contents of WM. In the keep-track task (Yntema, 1963), participants saw a list of 15 words drawn from multiple categories (e.g., animals, metals) presented one at a time on the screen. They were instructed to remember only the most recent word presented from each of two to five specified categories (displayed on the bottom of the screen). The dependent measure was the proportion of total words recalled across all 12 trials (four trials each of two, three, and four categories). In the letter memory task (Morris & Jones, 1990), participants saw a list of letters presented one at a time. After each letter, participants said aloud the three most recent letters that appeared (a rehearsal score). The dependent measure

²EF tasks elicit individual differences in both accuracy and RTs, and we include a mix of both in the task battery depending on what is most commonly done in previous studies. When RT is examined, there is a control condition that adjusts for the speed of performance (e.g., color-naming on the Stroop, single-task blocks for switching tasks) as there are ample individual difference in baseline processing speed. This is not typically done for accuracy tasks for which the control conditions (e.g., prosaccade) would be at ceiling.

³There was no baseline (prosaccade) condition in this task, so we used accuracy measures instead of unadjusted RTs. This dependent measure is identical to that used in previous work validating this model (Friedman et al., 2008; Miyake et al., 2000).

was the proportion of sets correctly rehearsed across 10 trials (four trials of length five letters, three trials of length seven and nine letters). Points were given for rehearsing only the correct letters in the correct serial order. In the spatial 2-back task (Friedman et al., 2008), participants saw an array of ten boxes scattered across the screen light up one at a time and pressed a button if the current box was the same as the one lit two trials before. The dependent measure was the total accuracy across four blocks of 24 trials.

Vocabulary.—Vocabulary was assessed in the same cognitive battery as verbal fluency (Defries et al., 1981). In the first subtest, participants had 3 minutes to answer up to 50 multiple-choice items consisting of stimulus words and four possible synonyms. In the second subtest, participants had 4 minutes to answer up to 25 more difficult items, this time with five response options. Dependent measures were the number of correct responses, adjusted for guessing. These two indicators of vocabulary were used to create a Vocabulary latent variable in all analyses.

Nonverbal cognitive ability.—Nonverbal cognitive ability was measured using the performance IQ score of the Wechsler Adult Intelligence Scale, third edition (WAIS; Wechsler, 1997). This score was based on perceptual organization and processing speed subtests including block design, matrix reasoning, picture completion, digit symbol-coding, and symbol search.

Sample 2: Middle-Aged Adults

Participants.—Analyses were based on 1290 male twins (363 full MZ pairs, 271 full DZ pairs, and 22 unpaired twins) in the VETSA who completed the wave 1 assessments ($M = 55.90$ years, $SD = 2.44$). VETSA participants were recruited from the Vietnam Era Twin Registry, and were randomly selected from Registry twins who participated in a previous study (Tsuang, Bar, Harley, & Lyons, 2001). All individuals served in the United States Military at some point between 1965 and 1975, but are representative of American men of their age with respect to health and lifestyle factors; nearly 80% did not serve in combat or Vietnam (Kremen et al., 2011; Kremen et al., 2006; Schoenborn & Heyman, 2009). Participants identified as white (90.8%), Asian (0.2%), black or African American (5.1%), American Indian (0.6%), more than one race (2.7%), or declined to answer (0.6%). Hispanic individuals comprised 3.0% of the sample.

Measures.

Verbal fluency.—Verbal fluency was assessed using the Delis-Kaplan Executive Function System (D-KEFS; Delis et al., 2001). Participants completed three phonemic fluency tasks (letters *F*, *A*, and *S*), followed by two semantic fluency tasks (*animals* and *boys' names*). Additionally, participants performed a category switching trial in which they alternated between naming fruits and pieces of furniture. To reduce the impact of the switching element of this task, dependent measures for each trial were the correct number of exemplars generated aloud within the 60-second response window, ignoring whether or not switches were made. Nevertheless, because the category switching trial involved some task switching, we allowed this trial to also be an indicator of Common EF in the current study. Its factor

loading was significant on the Common EF factor, but small in magnitude (see the appendix).

Executive function.—The EF battery in VETSA was based on six common neuropsychological measures of EF and has also been described in previous work (Gustavson, Panizzon, Franz, et al., 2018). The number of correct category switches was previously included in this model, but we omitted it from the current EF model to avoid biasing the association between the EF and verbal fluency factors. Removing this task did not affect the model, including the genetic and environmental estimates for either latent construct. Dependent measures for the Stroop task and Trail Making Test were multiplied by -1 in all analyses so that higher numbers always indicated better performance.

Inhibition was assessed using the Stroop task (Golden & Freshwater, 2002) and the AX-Continuous Performance Test (AX-CPT; Braver et al., 2001; Kremen et al., 2011). The dependent measure of the Stroop was a residualized score for the number of correct words identified during the color-word condition (naming colors of words printed in incongruent colors), adjusting for performance on the word condition (reading color words printed in black ink) and the color condition (naming the color of printed strings of *X*s). In the AX-CPT, participants saw a series of letters in the center of a computer screen. During 70% of the 150 trials, the letter *A* cue was followed by the letter *X* probe (*AX* trials), with only 10% of key trials where a non-*A* cue preceded the letter *X* probe (*BX* trials). The dependent measure was the arcsine-transformed signal detection index (the hit rate for *AX* trials minus the false alarm rate for *BX* trials).

Shifting was assessed with the Trail Making Test from the Delis-Kaplan Executive Function System (Delis et al., 2001). The dependent measure was the RT for switching trial (Condition 4) after residualizing the time on the letter sequencing and number sequencing trials (Conditions 2 and 3). WM span was assessed using the digit span and letter-number sequencing tasks from the Wechsler Memory Scale-III (Wechsler, 1997), and the reading span task (Daneman & Carpenter, 1980). In the digit span task, participants listened to strings of numbers and either repeated them directly (forward) or repeated them in reverse order (backward). In the letter-number sequencing task, participants re-sequenced letter-number strings of increasing size with numbers in ascending order, followed by the letters in alphabetical order. In both tasks, dependent measures were the total number of trials completed.⁴ In the reading span task, participants read multiple sentences aloud, and then recalled the last word of each sentence in order. The dependent measures was the total number of correct words recalled across all trials in the reading span (five trials each of length two, three, and four sentences).

Vocabulary.—Vocabulary was assessed using the vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI-V; Wechsler, 1999) and the timed multiple-choice vocabulary subtest of the Armed Forces Qualification Test (AFQT-V; Bayroff & Anderson,

⁴We followed standard administration in which subjects complete two (digit span) or three (letter-number sequencing) trials of each span length (starting with two items), increasing in difficulty until both trials of a given span length were incorrect. For the digit span only, subjects were administered two additional span lengths before discontinuing the test. The maximum span length was seven for letter-number sequencing, ten for digit span forward, and nine for digit span backwards.

1963; Lyons et al., 2017). Based on modification indices, we also included a factor loading from the Vocabulary latent variable to the reading span task, likely reflecting the additional reading demands in the task compared to the other span measures. However, excluding this additional factor loading has no impact on the main conclusions.

Nonverbal cognitive ability.—Nonverbal cognitive ability was assessed using the matrix reasoning subtest of the WASI (WASI-M; Wechsler, 1999) and the average of the scores for the two primarily nonverbal components the AFQT: box folding (visuospatial organization) and knowledge and reasoning about tools and mechanical relations (AFQT-G; Bayroff & Anderson, 1963; Lyons et al., 2017). These two measures were combined to create a latent variable.

Data Analysis

In both samples, all measures were adjusted for age by creating residualized scores after accounting for age at that wave of assessment. When creating these residualized scores, we also adjusted for sex in the adolescent sample.

Analyses were conducted using Mplus for phenotypic analyses (Muthén & Muthén, 2010–2012) and the structural equation modeling package OpenMx in R for genetic analyses (Neale et al., 2016). Both programs account for missing observations using full-information maximum likelihood. Model fit was determined based on chi-square tests (χ^2), the root mean error of approximation (RMSEA), and the Comparative Fit Index (CFI). Models with the best fit had χ^2 values less than two times the degrees of freedom, RMSEA values less than .06, and CFI values greater than .95 (Hu & Bentler, 1998). Genetic models were also compared to a full genetic Cholesky decomposition using χ^2 difference tests. Significance of individual parameter estimates was established using standard error-based 95% confidence intervals (CIs) in phenotypic analyses and likelihood-based CIs in genetic analyses (Neale, Heath, Hewitt, Eaves, & Fulker, 1989). 95% CIs are reported in the main text, but we also confirmed that $p < .05$ values reported in the tables were accurate with χ^2 difference tests. Phenotypic analyses accounted for the clustering by family in the data with Mplus's "type=complex" command, which yields standard errors and a χ^2 statistic that are adequately adjusted for non-independence of twin data (Rebollo, de Moor, Dolan, & Boomsma, 2006; Satorra & Bentler, 2001).

Genetic analyses.—Genetically informed models were based on standard assumptions in twin designs (Neale & Cardon, 1992), in which the variance of a phenotype can be separated into proportions attributable to additive genetic influences (A), common environmental influences (C), and nonshared environmental influences (E). Additive genetic influences (A) are correlated at 1.0 in MZ twin pairs and 0.5 in DZ twin pairs because MZ twins share 100% of their alleles identical-by-descent and DZ twins share, on average, 50% of their alleles identical-by-descent. Common/shared environmental influences (C) correlate 1.0 for both MZ and DZ twins because they are environmental factors that make twins more similar to one another. Nonshared environmental influences (E) do not correlate. These are defined as environmental factors that make twins different from each other. We also assume equal means and variances within pairs and across zygosity. These standard assumptions for

univariate twin analyses extend to the multivariate analyses described here, including situations where phenotypic correlations are decomposed into their genetic (r_g), shared environmental (r_c), and nonshared environmental components (r_e). As in other structural equation models, latent genetic/environmental variance components were not measured directly. Rather, these assumptions were imposed on the model and the maximum likelihood estimator converged on the best-fitting solution.

The latent variable models and univariate heritability's for most measures were validated in earlier work (Friedman et al., 2008; Gustavson, Panizzon, Franz, et al., 2018; Kremen, Moore, Franz, Panizzon, & Lyons, 2014). Therefore, for each sample, we present only the common pathway models of EFs alone, verbal fluency alone, and the combined model of EFs and fluency together. In genetic models with both EFs and fluency, we used a Cholesky decomposition to examine their overlapping and unique genetic/environmental variance components. Figure 1D displays this Cholesky decomposition. Genetic and environmental correlations between variables were algebraically computed from this Cholesky decomposition.⁵ We did not estimate shared environmental correlations in these models because shared environmental influences were estimated at $c^2=.00$ on EFs in adolescents (Friedman et al., 2008) and on verbal fluency in middle-aged adults (Gustavson, Panizzon, Elman, Beck, et al., 2018) in these data. Additionally, we did not test for nonshared environmental correlations with the Updating-Specific factor in adolescents as they were also estimated at $e^2=.00$.

Results

Descriptive Statistics

Descriptive statistics for all measures in both samples are displayed in Table 2, and phenotypic correlations among all measures are displayed in the supplement (Tables S1 and S2).

Models of Executive Functions and Verbal Fluency Alone

All genetic/environmental models of EFs (Figure 2) and verbal fluency (Figure 3) fit well (see figure captions for model fit). In these models, percent variance explained in individual EF/fluency tasks (shown in rectangles) by their respective latent factors can be computed by squaring the factor loadings on the latent variables (shown in ovals). Similarly, variance in latent EF/fluency factors explained by genetic (A), shared environmental (C), and nonshared environmental influences (E) can be computed by squaring their factor loadings.

Executive functions.—In adolescents (Figure 2a), genetic influences explained almost all the variance in the Common EF, $a^2=.97$, 95% CI [.63, 1.0], Updating-Specific, $a^2=1.0$, 95% CI [.63, 1.0], and Shifting-Specific factors, $a^2=.79$, 95% CI [.38, .96]. Nonshared

⁵In these models, we did not allow cross-paths between orthogonal latent variables. However, because fluency factors were included last in the Cholesky order, there were implied correlations between the two fluency factors through their associations with EFs (and vocabulary or nonverbal cognitive ability in the full model). Importantly, these implied correlations were small, and including fluency last in the ordering was necessary to test the significance of their unique genetic/environmental variance components.

environmental influences were only significant for the Shifting-Specific factor, $e^2=.21$, 95% CI [.04, .41].

In middle-aged adults (Figure 2b), genetic influences on the Common EF factor were somewhat smaller than in adolescents, $a^2=.41$, 95% CI [.07, .80], but were substantial for the WM-Specific factor, $a^2=.85$, 95% CI [.51, .94]. The smaller heritability for the Common EF factor in middle-aged adults was due to a larger contribution of shared environmental influences, $c^2=.47$, 95% CI [.11, .77]. Nonshared environmental influences were significant for both Common EF, $e^2=.12$, 95% CI [.03, .22], and WM-Specific factors, $e^2=.15$, 95% CI [.06, .25].

Verbal fluency.—In adolescents (Figure 3a), the General Fluency factor was mostly explained by genetic influences, $a^2=.72$, 95% CI [.35, .94], with some nonshared environmental influences, $e^2=.14$, 95% CI [.05, .25]. The Semantic-Specific factor was also explained mostly by genetic influences, $a^2=.80$, 95% CI [.28, .97], with some nonshared environmental influences, $e^2=.18$, 95% CI [.02, .36].

In middle-aged adults (Figure 3b), genetic influences accounted for most of the variance in the General Fluency factor, $a^2=.76$, 95% CI [.66, .82], with some nonshared environmental influences, $e^2=.24$, 95% CI [.18, .30]. Genetic influences also accounted for most of the variance in the Semantic-Specific factor, $a^2=.65$, 95% CI [.26, .85], with significant nonshared environmental influences, $e^2=.35$, 95% CI [.15, .57].

Phenotypic Models of Executive Function and Verbal Fluency

We first examined the phenotypic correlations between EFs, verbal fluency, vocabulary, and nonverbal cognitive ability in each sample. The latent variable correlations between all constructs are displayed in Table 3. Next, we conducted multiple regression analyses to examine whether associations between EFs and fluency factors remained significant after controlling for vocabulary and nonverbal cognitive ability. Standardized regression coefficients are displayed in Table 4. Factor loadings and model fit for correlational and regression models were identical. Factor loadings are displayed in the supplement (Table S3) alongside path diagrams (Figure S1).

Adolescents.—As shown in Table 3, the General Fluency factor was significantly correlated with the Common EF, Updating-Specific, and Shifting-Specific factors. Regression analyses in Table 4 revealed that these associations all remained statistically significant even when controlling for vocabulary and nonverbal cognitive ability (see supplement for additional analyses involving processing speed). In the regression model, the General Fluency factor was positively associated with the Common EF, Updating-Specific, and Vocabulary factors, and negatively associated with the Shifting-Specific factor. Although the General Fluency factor was also positively correlated with nonverbal cognitive ability, this association was nonsignificant in the regression model. In total, half of the variance in the General Fluency factor was captured by the other abilities, $R^2=.50$.

The Semantic-Specific factor was not associated with the Common EF factor, but it was positively correlated with both the Updating-Specific and Shifting-Specific factors. In the

regression model, the Semantic-Specific factor remained positively associated with the Shifting-Specific factor, the Vocabulary factor, and the measure of nonverbal cognitive ability. However, the correlation with the Updating-Specific factor was no longer significant. In total, almost half of the variance in the Semantic-Specific factor was captured by all the cognitive variables ($R^2=.43$).

Middle-aged adults.—In middle-aged adults, the General Fluency factor was positively correlated with both the Common EF and WM-Specific factors, as well as the vocabulary factor. Just as in adolescents, these associations remained significant in the regression model. Also like adolescents, the positive correlation between the General Fluency and nonverbal cognitive ability factors was no longer significant after controlling for EFs and vocabulary. In total, about a third of the variance in the General Fluency factor was explained ($R^2=.33$).

The Semantic-Specific factor was weakly positively associated with the Common EF factor in middle-aged adults, but this association was nonsignificant in the regression analyses. There were no associations between the Semantic-Specific and WM-Specific factors in the correlational or regression models. Although the Semantic-Specific factor was positively correlated with both the vocabulary and nonverbal cognitive ability factors to roughly the same degree as in adolescents, only the association with vocabulary reached statistical significance in the regression model. In total, only about 11% of the variance in the Semantic-Specific factor was explained ($R^2=.11$).

Summary of phenotypic analyses.—Phenotypic results supported the third scenario. In both samples, the General Fluency factor was moderately correlated with the Common EF factor, but it was also associated with each of the domain-specific EF factors. Furthermore, at least half of the variance in the General Fluency factor was independent of EFs, vocabulary, and nonverbal cognitive ability. As expected, the Semantic-Specific factor was not associated with Common EF (with some caveats), but it was associated with the Shifting-Specific factor in adolescents. In both samples, much of the variance in the Semantic-Specific factor was unique.

Genetic Models of Executive Function and Verbal Fluency

Next, we combined genetic/environmental models of EFs and verbal fluency from each sample, displayed in Table 5. Factor loadings on individual tasks and genetic/environmental variance components on latent factors are not displayed, but were nearly identical to those depicted in Figures 2 and 3.

Adolescents.—As shown in Table 5 (top), most of the phenotypic associations between the EF and fluency factors were explained by genetic influences. Genetic influences explained most of the variance in the Common EF and General Fluency factors, so it was not surprising that 80% of their phenotypic correlation was explained by a significant genetic correlation ($r_g=.37$). There was also a large nonshared environmental correlation between the General Fluency and Common EF factors ($r_e=.84$), but it explained only 20% of the phenotypic correlation given the weaker contribution of environmental influences on both constructs. The Updating-Specific factor was entirely captured by genetic influences, so they

explained the entire phenotypic correlation with the General Fluency factor. The negative association between the General Fluency and Shifting-Specific factors was explained roughly equally by genetic (45%) and nonshared environmental (55%) influences, but only the nonshared environmental correlation was significant.

For the Semantic-Specific factor, there was a significant genetic correlation with the Updating-Specific factor ($r_g=.38$). Additionally, the Semantic-Specific factor was genetically correlated with the Shifting-Specific factor ($r_g=.30$), explaining 62% of the phenotypic correlation, but there was also a significant nonshared environmental correlation ($r_e=.58$).

Importantly, there were substantial unique genetic influences on both fluency factors after accounting for their associations with EFs. Genetic influences independent of EFs accounted for 45% of the variance in the General Fluency factor (compared to 81%, the total genetic variance), 95% CI [.15, .58], suggesting that over half of the total genetic variance on the General Fluency factor was independent of EFs. Similarly, 50% of the variance in the Semantic-Specific factor was explained by genetic influences independent of EF factors (compared to 66% total), 95% CI [.29, .69]. In contrast, there were no significant unique environmental influences on either factor after accounting for EF factors (0% for the General Fluency factor, 95% CI [.00, .18], 15% for the Semantic-Specific factor, 95% CI [.00, .35]).

Middle-aged adults.—Genetic influences explained 91% of the phenotypic correlation between the General Fluency and Common EF factors ($r_g=.71$). The nonshared environmental correlation between the General Fluency and Common EF factors was nonsignificant. There was also a significant genetic correlation between the General Fluency and WM-Specific factors ($r_g=.40$), explaining 94% of their phenotypic correlation. Again, the nonshared environmental correlation was not significant. Consistent with the lack of phenotypic correlation, the Semantic-Specific factor was not associated with either the Common EF or WM-Specific factors at the genetic or nonshared environmental levels.

After accounting for the overlap with EFs, unique genetic influences on the General Fluency factor accounted for 25% of the variance (compared to 76% total), but they were nonsignificant, 95% CI [.00, .45]. Nonshared environmental influences on the General Fluency factor independent of other cognitive abilities were comparable in magnitude and significant, $e^2=.23$, 95% CI [.15, .30]. There were also unique genetic (46%, 95% CI [.11, .74]) and nonshared environmental influences (33%, 95% CI [.03, .54]) on the Semantic-Specific factor beyond those explained by EF factors.

Summary of genetic analyses.—Almost all the association between the General Fluency and Common EF factors was explained by genetic influences, with some evidence for a nonshared environmental correlation in adolescents only. Genetic influences also accounted for the associations between the General Fluency and Updating-Specific (adolescents) or WM-Specific factors (middle-aged adults). Although we were only able to assess a Shifting-Specific factor in adolescents, its association with fluency was relatively equally explained by genetic and nonshared environmental influences both with respect to its negative correlation with the General Fluency factor and its positive correlation with the Semantic-Specific factor. Importantly, there were substantial genetic influences on both

fluency factors even after accounting for EFs (and other cognitive abilities; see Supplemental Results).

Discussion

Our results showed that fluency reflects many sources of EF variance, but also vocabulary and two sets of unique genetic influences. These results were largely consistent across age and samples representing two developmental periods.

Implications for Verbal Fluency

If verbal fluency ability is akin to shifting or WM updating, we would expect to observe associations between the General Fluency and Common EF factors, but no associations between the General Fluency factor and the domain-specific EF factors. However, this was not the case. The General Fluency factor was associated with the Common EF factor, but also associated with the Updating-Specific and Shifting-Specific factors in adolescents, and with the WM-Specific factor in middle aged adults. To the extent that verbal fluency draws on both Common EF ability and an assortment of domain-specific EF abilities, fluency may be viewed as a combination of other EF abilities. However, because there was evidence for unique variance in the General Fluency factor beyond that explained by EF factors, vocabulary, and nonverbal cognitive ability, it is clearly capturing some unique sources of variance as well.

These results are also relevant given the widespread use of fluency measures, especially in the context of neuropsychiatric conditions such as Alzheimer's disease, Parkinson's disease, schizophrenia, and depression (Henry & Crawford, 2004, 2005, 2005; Henry et al., 2004; Ho, Nation, & Alzheimer's Disease Neuroimaging, 2018). These findings suggest that measures of verbal fluency may be good at capturing global EF impairments, as they reflect multiple unique EF-relevant variance components. However, this combination of multiple abilities makes determining the underlying mechanisms driving these associations more difficult. Therefore, one must consider multiple mechanisms when interpreting research findings using fluency tasks, especially in designs that do not have measures of EFs and vocabulary for comparison.

The genetically informative nature of the samples analyzed here enabled us to better characterize the associations between fluency and EF abilities. For example, the phenotypic associations between fluency and EF factors were driven by genetic influences at both ages. Environmental correlations were significant in adolescents (between General Fluency and Common EF and between the Semantic-Specific and Shifting-Specific factors), but not middle-aged adults. These correlations were explained by nonshared environmental influences (that make twins different) as opposed to shared environmental influences (that make twins similar). It will be useful to pinpoint the environmental factors that account for these associations. However, just as genetic correlations reflect the combined effects of many independent genetic influences, the nonshared environmental correlation may reflect a multitude of factors, each of which only accounts for a small fraction of the variance (Friedman et al., 2016; Plomin, Asbury, & Dunn, 2001).

In addition, there were unique genetic influences on the fluency factors beyond those explained by other EF abilities. Some of these unique genetic influences may reflect other cognitive abilities not measured here (e.g., episodic memory, processing speed), but we should have captured the most relevant cognitive abilities by controlling for vocabulary and nonverbal cognitive ability. Indeed, although speed is correlated with almost all of the cognitive abilities assessed here (Friedman et al., 2008; Gustavson, Panizzon, Franz, et al., 2018; Unsworth et al., 2011), additional analyses described in the supplement (including measures of perceptual speed in adolescents or reaction time in middle-aged adults) demonstrated that all associations in the multiple regression (Table 4) remained significant, with minimal increases in R^2 of the models.

Future work could also examine these associations in the context of other dependent measures of fluency. For instance, the number of correct words generated may reflect the contribution of different sub-processes or strategies (Troyer & Moscovitch, 2006). Although we were not able to examine other dependent measures here, the associations between fluency, EFs, and vocabulary latent factors may differ across them (Unsworth et al., 2011). It will be interesting to examine these associations in the context of the unity/diversity model and quantify whether alternate measures capture unique genetic influences from one other (or from the total words generated).

Implications for Executive Functions

Inhibition, shifting, and WM updating are the most widely studied EF abilities, but it is a common misperception that these are the only three types of EFs (Friedman & Miyake, 2017). Research on verbal fluency has developed somewhat independently from other EFs. Our findings suggest that verbal fluency ability draws on the same set of Common EF variance as inhibition, shifting, and updating abilities. However, we are not proposing that fluency should necessarily be incorporated into the unity/diversity model. Fluency tasks draw on multiple EF processes (Common EF, Updating-Specific, Shifting-Specific abilities) as well as vocabulary and at least two sources of unique influences (General Fluency and Semantic-Specific abilities). Thus, it may be best to include fluency factors in regression models with EF and vocabulary factors when examining associations with other constructs (e.g., rather than including factor loadings from individual fluency tasks to as many as five EF factors). This approach has been adopted in some recent studies on the association between fluency and other cognitive phenotypes (Hedden, Lautenschlager, & Park, 2005; Hedden & Yoon, 2006; Unsworth et al., 2011), but has not yet been widely applied in other domains.

These results are relevant to our understanding of Updating-Specific and Shifting-Specific abilities, which have remained somewhat unclear (Friedman & Miyake, 2017). The negative association between the Shifting-Specific and General Fluency factors is consistent with other findings that Shifting-Specific ability is negatively correlated with other positive traits such as intelligence (Friedman et al., 2008), and supports the idea that it may trade-off with the strength of goal activation, thought to be represented by the Common EF factor. For example, Herd et al. (2014) developed a computational model demonstrating that Shifting-Specific ability may reflect uncontrolled, automatic persistence of goal representations, and

that strong goals (i.e., as would be seen with higher Common EF) persist longer because they are more active. However, the moderate positive correlation with the Semantic-Specific factor is unique in that it is one of the first positive associations observed. Further studies are needed, but these results provide some initial evidence for the beneficial aspects of Shifting-Specific abilities (i.e., abilities that are unique to shifting and not part of Common EF).

These results are also relevant to our understanding of Updating-Specific ability. Although it is also unclear, Updating-Specific ability may reflect the effective gating of information in the basal ganglia or controlled retrieval from long-term memory (Miyake & Friedman, 2012). The latter possibility is consistent with the positive associations between the General Fluency factor and the variance unique to WM processes in both age groups (i.e., Updating-Specific and WM-Specific factors), which likely reflect the search through semantic long-term memory. Regression analyses revealed no significant associations between these WM factors and the Semantic-Specific factor after controlling for vocabulary and nonverbal cognitive ability. This is interesting because semantic fluency requires search through real-world semantic categories (e.g., animals) as opposed to lexical categories (e.g., *S* words), but no extra WM processes appear to be involved in semantic fluency that are not explained by variance shared with phonemic fluency (i.e., General Fluency).

Examining these data at different ages also grants insights into the generalizability of our findings and potential age differences in these associations. The associations between the General Fluency and Common EF factors were nearly identical in adolescents and middle-aged adults, and driven by genetic influences, supporting the idea that there appear to be few changes in the structure and genetic influences underlying EFs and their covariance with fluency across adulthood. The only task in common between samples was the Stroop, suggesting that these findings generalize well across the different measures and administration methods, especially at the latent variable level. Furthermore, each of the fluency and EFs factors decline at different rates across middle age (Gustavson, Panizzon, Elman, Beck, et al., 2018; Gustavson, Panizzon, Elman, Franz, et al., 2018), yet vocabulary continues to improve throughout the lifespan (Park & Reuter-Lorenz, 2009). Despite these differences, their pattern of covariance in middle-aged adults remains remarkably similar to those in adolescents, although longitudinal studies will be necessary to examine the dynamics of these associations in more detail.

Nonverbal Cognitive Ability and Vocabulary

We also examined whether other cognitive abilities were relevant to these associations. The Common EF factor is moderately-to-strongly genetically correlated with measures of intelligence or general cognitive ability (Engelhardt et al., 2016; Friedman et al., 2008; Gustavson, Panizzon, Franz, et al., 2018), but these findings further demonstrate that these constructs are not the same. Measures of nonverbal general cognitive ability were correlated with all the constructs in the study, but multiple regression analyses indicated that these associations with fluency factors were nonsignificant after controlling for EFs and vocabulary, with one exception for the Semantic-Specific factor in adolescents. Thus, despite the moderate genetic overlap between EFs and nonverbal cognitive ability, it is the EF variance independent of nonverbal cognitive ability that is most relevant to verbal fluency.

In contrast, vocabulary was positively associated with both fluency factors, and this association was similar in magnitude as those for EFs in the regression analyses. These findings suggest that fluency tasks are not pure measures of EFs; vocabulary plays a considerable role. This is consistent with work by Whiteside et al. (2016) who observed that fluency was more strongly related to vocabulary than EFs in a heterogeneous sample of outpatients, though they did not examine these associations at the level of latent variables. Moreover, results indicated that these associations with vocabulary were driven almost exclusively by genetic influences for General Fluency ability (99–100%) but by environmental influences for Semantic-Specific ability (66% in adolescents, 98% in middle-aged adults). Although power to decompose these associations were low in the full genetic model (described briefly in the supplemental results), they suggest that environmental influences on vocabulary play a larger role in semantic than phonemic fluency.

Limitations

There were some differences in the samples and measures that affect our ability to draw stronger conclusions about the correspondence of these results. First, the sample of adolescents comprised both men and women, but the middle-aged sample comprised only men. Second, the domain-specific EF factors differed in each sample (e.g., Shifting-Specific and Updating-Specific factors in adolescents, WM-Specific factor in middle-aged adults), restricting the generalizability of their results. Ideally, we would have conducted this study using the same sample of individuals at both timepoints, and using the same measures of EFs and fluency, though such studies rarely exist. Nevertheless, the strong correspondence between the results of both samples suggest that these findings do generalize across assessment and age and will be valuable to our growing understanding of the multiple cognitive domains thought to be captured by EFs.

We did not have the power to distinguish between additive and dominant genetic influences in this study (Martin, Eaves, Kearsy, & Davies, 1978), or examine epigenetic influences, gene/environment correlations, or gene-by-environment interactions, which can only be examined using much larger samples. Finally, the associations between verbal fluency, EFs, and other cognitive abilities may differ in clinical samples. However, because both samples were representative of the general populations the estimates of heritability were probably not biased (e.g., by an overly healthy sample).

Summary and Conclusions

Fluency has historically been described as an EF, but it has been difficult to interpret results due to the lack of elucidation of the correspondence between fluency and other well-studied EF processes (inhibition, shifting, and WM updating) and their commonality (Common EF). This study was the first to examine the genetic/environmental overlap between EFs and verbal fluency using the recent unity/diversity model of EFs that highlights Common EF and domain-specific variance. Verbal fluency factors were associated with EF factors in part through associations with the Common EF factor, but also through other EFs (e.g., the WM-Specific factor). A second fluency factor representing Semantic-Specific fluency was not correlated with the Common EF factor at either adolescents or middle-aged adults, but it was associated with Shifting-Specific and vocabulary factors. It will be important to consider the

multi-faceted nature of fluency in future work, as performance on semantic fluency tasks seems to be captured by at least five different sources of variance, each of which has some unique genetic and environmental influences. Given the relevance of verbal fluency and EFs to social, cognitive, and health phenotypes across the lifespan, it will be useful to examine their associations in the context of large, genetically-informative models which are well suited to isolate the variance components most relevant to these associations.

Context of the Research

Our work has highlighted the commonality and specificity of various processes thought to rely on executive function. Studies typically focus on measures of inhibition, working memory updating, and task-set shifting, but the relation of verbal fluency to other executive processes has long been of interest. Fluency is sometimes considered an indicator of executive function, but also sometimes considered a measure of language ability. This work shows that variance in verbal fluency reflects a wide array of cognitive abilities, including both general executive function processes, processes specific to some executive constructs and not others (updating-specific and shifting-specific abilities), and other non-executive cognitive abilities (vocabulary). These results are consistent with the utility of verbal fluency tasks as neuropsychological measures sensitive to broad range of cognitive impairments. However, our findings suggest that measures of verbal fluency may not be highly useful as a research tool if the goal is to isolate specific mechanisms, as they blend together many sources of cognitive individual differences. Therefore, although these results suggest that verbal fluency cannot be neatly incorporated into models of executive function (such as the unity/diversity model we highlight here), there are at least two sets of genetic influences underlying fluency tasks that could not be attributed to executive functions or other cognitive abilities. These unique sets of variance in verbal fluency, especially the variance unique to semantic fluency, may be particularly relevant to future studies of cognition, health, and aging.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by Grants AG050595, AG018386, AG018384, AG022381, AG047903, AG046938, MH063207, and AG046938 from the National Institutes of Health.

The content of this manuscript is the responsibility of the authors and does not represent official views of NIA/NIH, or the Veterans' Administration. Numerous organizations provided invaluable assistance in the conduct of the VET Registry, including: U.S. Department of Veterans Affairs, Department of Defense; National Personnel Records Center, National Archives and Records Administration; Internal Revenue Service; National Opinion Research Center; National Research Council, National Academy of Sciences; the Institute for Survey Research, Temple University. The authors gratefully acknowledge the continued cooperation of the twins and the efforts of many staff members

References

Alvarez JA, & Emory E (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, 16, 17–42. doi:10.1007/s11065-006-9002-x [PubMed: 16794878]

- Bayroff AG, & Anderson AA (1963). Development of Literacy Screening Scales for AFQT 7 and 8 Failures. Washington DC:
- Braver TS, Barch DM, Keys BA, Carter CS, Cohen JD, Kaye JA, ... Reed BR (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, 130, 746–763. doi: 10.1037/0096-3445.130.4.746 [PubMed: 11757878]
- Daneman M, & Carpenter PA (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466. doi:10.1016/S0022-5371(80)90312-6
- Defries JC, Plomin R, Vandenberg SG, & Kuse AR (1981). Parent-offspring resemblance for cognitive-abilities in the Colorado Adoption Project: Biological, adoptive, and control parents and one-year-old children. *Intelligence*, 5, 245–277. doi:10.1016/S0160-2896(81)80012-8
- DeFries JC, Vandenberg SG, McClearn GE, Kuse AR, Wilson JR, Ashton GC, & Johnson RC (1974). Near identity of cognitive structure in two ethnic groups. *Science*, 183, 338–339. doi:10.1126/science.183.4122.338 [PubMed: 4809078]
- Delis DC, Kaplan E, & Kramer JH (2001). Delis-Kaplan executive function system (D-KEFS): Psychological Corporation.
- Diamond A (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. doi:10.1146/annurev-psych-113011-143750
- Ekstrom RB, French JW, Harman HH, & Derman DD (1976). Kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service.
- Engelhardt LE, Briley DA, Mann FD, Harden KP, & Tucker-Drob EM (2015). Genes unite executive functions in childhood. *Psychological Science*, 26, 1151–1163. doi:10.1177/0956797615577209 [PubMed: 26246520]
- Engelhardt LE, Mann FD, Briley DA, Church JA, Harden KP, & Tucker-Drob EM (2016). Strong genetic overlap between executive functions and intelligence. *Journal of Experimental Psychology: General*, 145, 1141–1159. doi:10.1037/xge0000195 [PubMed: 27359131]
- Fisk JE, & Sharp CA (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, 26, 874–890. doi: 10.1080/13803390490510680 [PubMed: 15742539]
- Frank MJ, Loughry B, & O'Reilly RC (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive Affective & Behavioral Neuroscience*, 1, 137–160. doi:10.3758/Cabn.1.2.137
- Friedman NP, & Hewitt JK (2013). Do working memory span and updating tasks measure the same thing? A twin study. Paper presented at the Behavioral Genetics Annual Meeting.
- Friedman NP, & Miyake A (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86, 186–204. doi:10.1016/j.cortex.2016.04.023 [PubMed: 27251123]
- Friedman NP, Miyake A, Altamirano LJ, Corley RP, Young SE, Rhea SA, & Hewitt JK (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology*, 52, 326–340. doi:10.1037/dev0000075 [PubMed: 26619323]
- Friedman NP, Miyake A, Robinson JL, & Hewitt JK (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, 47, 1410–1430. doi:10.1037/a0023750 [PubMed: 21668099]
- Friedman NP, Miyake A, Young SE, Defries JC, Corley RP, & Hewitt JK (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, 137, 201–225. doi:10.1037/0096-3445.137.2.201 [PubMed: 18473654]
- Golden CJ, & Freshwater SM (2002). The Stroop color and word test: A manual for clinical and experimental uses[adult version]: Stoelting.
- Goschke T (2000). Intentional reconfiguration and involuntary persistence in task set switching In Monsell S & Driver J (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 331–355). Cambridge, MA: MIT Press.

- Gustavson DE, Panizzon MS, Elman JA, Beck A, Franz CE, Reynolds CA, ... Kremen WS (2018). Genetic and environmental influences on verbal fluency in middle age: A longitudinal twin study. *Behavior Genetics*, 48, 361–373. doi:10.1007/s10519-018-9910-6 [PubMed: 29922985]
- Gustavson DE, Panizzon MS, Elman JA, Franz CE, Reynolds CA, Jacobson KC, ... Kremen WS (2018). Stability of genetic and environmental influences on executive functions in midlife. *Psychology and Aging*, 33, 219–231. doi:10.1037/pag0000230 [PubMed: 29658745]
- Gustavson DE, Panizzon MS, Franz CE, Friedman NP, Reynolds CA, Jacobson KC, ... Kremen WS (2018). Genetic and environmental architecture of executive functions in midlife. *Neuropsychology*, 32, 18–30. doi:10.1037/neu0000389 [PubMed: 28857599]
- Hedden T, Lautenschlager G, & Park DC (2005). Contributions of processing ability and knowledge to verbal memory tasks across the adult life-span. *The Quarterly Journal of Experimental Psychology*, 58, 169–190. doi:10.1080/02724980443000179 [PubMed: 15881297]
- Hedden T, & Yoon C (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology*, 20, 511–528. doi:10.1037/0894-4105.20.5.511 [PubMed: 16938014]
- Henry JD, & Crawford JR (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, 18, 284–295. doi:10.1037/0894-4105.18.2.284 [PubMed: 15099151]
- Henry JD, & Crawford JR (2004). Verbal fluency deficits in Parkinson's disease: A meta-analysis. *Journal of the International Neuropsychological Society*, 10, 608–622. doi:10.1017/S1355617704104141 [PubMed: 15327739]
- Henry JD, & Crawford JR (2005). A meta-analytic review of verbal fluency deficits in depression. *Journal of Clinical and Experimental Neuropsychology*, 27, 78–101. doi:10.1080/138033990513654 [PubMed: 15814444]
- Henry JD, & Crawford JR (2005). A meta-analytic review of verbal fluency deficits in schizophrenia relative to other neurocognitive deficits. *Cognitive Neuropsychiatry*, 10, 1–33. doi:10.1080/13546800344000309 [PubMed: 16571449]
- Henry JD, Crawford JR, & Phillips LH (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, 42, 1212–1222. doi:10.1016/j.neuropsychologia.2004.02.001 [PubMed: 15178173]
- Herd SA, O'Reilly RC, Hazy TE, Chatham CH, Brant AM, & Friedman NP (2014). A neural network model of individual differences in task switching abilities. *Neuropsychologia*, 62, 375–389. doi:10.1016/j.neuropsychologia.2014.04.014 [PubMed: 24791709]
- Ho JK, Nation DA, & Alzheimer's Disease Neuroimaging I (2018). Neuropsychological profiles and trajectories in preclinical Alzheimer's Disease. *Journal of the International Neuropsychological Society*, 24, 693–702. doi:10.1017/S135561771800022X [PubMed: 29706146]
- Hu LT, & Bentler PM (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424–453. doi:10.1037/1082-989x.3.4.424
- Ito TA, Friedman NP, Bartholow BD, Correll J, Loersch C, Altamirano LJ, & Miyake A (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, 108, 187–218. doi:10.1037/a0038557 [PubMed: 25603372]
- Kremen WS, Jacobson KC, Panizzon MS, Xian H, Eaves LJ, Eisen SA, ... Lyons MJ (2009). Factor structure of planning and problem-solving: A behavioral genetic analysis of the Tower of London task in middle-aged twins. *Behavior Genetics*, 39, 133–144. doi:10.1007/s10519-008-9242-z [PubMed: 19085096]
- Kremen WS, Moore CS, Franz CE, Panizzon MS, & Lyons MJ (2014). Cognition in middle adulthood. In Finkel D & Reynolds CA (Eds.), *Behavior genetics of cognition across the lifespan* (pp. 105–134): Springer New York.
- Kremen WS, Panizzon MS, Xian H, Barch DM, Franz CE, Grant MD, ... Lyons MJ (2011). Genetic architecture of context processing in late middle age: More than one underlying mechanism. *Psychology and Aging*, 26, 852–863. doi:10.1037/a0025098 [PubMed: 21875218]

- Kremen WS, Thompson-Brenner H, Leung YM, Grant MD, Franz CE, Eisen SA, ... Lyons MJ (2006). Genes, environment, and time: The Vietnam Era Twin Study of Aging (VETSA). *Twin Research and Human Genetics*, 9, 1009–1022. doi:10.1375/183242706779462750 [PubMed: 17254445]
- Logan GD (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm In D D & Carr TH (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.
- Lyons MJ, Panizzon MS, Liu W, McKenzie R, Bluestone NJ, Grant MD, ... Xian H (2017). A longitudinal twin study of general cognitive ability over four decades. *Developmental Psychology*, 53, 1170–1177. doi:10.1037/dev0000303 [PubMed: 28358535]
- Lyons MJ, York TP, Franz CE, Grant MD, Eaves LJ, Jacobson KC, ... Kremen WS (2009). Genes determine stability and the environment determines change in cognitive ability during 35 years of adulthood. *Psychological Science*, 20, 1146–1152. doi:10.1111/j.1467-9280.2009.02425.x [PubMed: 19686293]
- Martin NG, Eaves LJ, Kearsy MJ, & Davies P (1978). The power of the classical twin study. *Heredity* (Edinb), 40, 97–116. doi:10.1038/hdy.1978.10 [PubMed: 272366]
- Mayr U, & Kliegl R (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1124–1140.
- Miyake A, Emerson MJ, Padilla F, & Ahn JC (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, 115, 123–142. doi:10.1016/j.actpsy.2003.12.004 [PubMed: 14962397]
- Miyake A, & Friedman NP (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21, 8–14. doi:10.1177/0963721411429458 [PubMed: 22773897]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100. doi:10.1006/cogp.1999.0734 [PubMed: 10945922]
- Morris N, & Jones DM (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, 81, 111–121. doi:10.1111/j.2044-8295.1990.tb02349.x
- Muthén LK, & Muthén BO (2010-2012). *Mplus Version 7 user's guide*. Los Angeles, CA: Muthén & Muthén.
- Neale MC, & Cardon LR (1992). *Methodology for genetic studies of twins and families*. Dordrecht:: Kluwer Academic Publishers.
- Neale MC, Heath AC, Hewitt JK, Eaves LJ, & Fulker DW (1989). Fitting genetic models with LISREL: hypothesis testing. *Behavior Genetics*, 19, 37–49. doi:10.1007/BF01065882 [PubMed: 2712812]
- Neale MC, Hunter MD, Pritikin JN, Zahery M, Brick TR, Kirkpatrick RM, ... Boker SM (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, 81, 535–549. doi:10.1007/s11336-014-9435-8 [PubMed: 25622929]
- Orme DR, Brehm W, & Ree MJ (2001). Armed forces qualification test as a measure of premorbid intelligence. *Military Psychology*, 13, 187–197. doi:10.1207/S15327876mp1304_1
- Park DC, & Reuter-Lorenz P (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60, 173–196. doi:10.1146/annurev.psych.59.103006.093656
- Phillips LH, Bull R, Adams E, & Fraser L (2002). Positive mood and executive function: evidence from stroop and fluency tasks. *Emotion*, 2, 12–22. doi:10.1037/1528-3542.2.1.12 [PubMed: 12899364]
- Plomin R, Asbury K, & Dunn J (2001). Why are children in the same family so different? Nonshared environment a decade later. *The Canadian Journal of Psychiatry*, 46, 225–233. doi:10.1177/070674370104600302 [PubMed: 11320676]
- Rebollo I, de Moor MHM, Dolan CV, & Boomsma DI (2006). Phenotypic factor analysis of family data: Correction of the bias due to dependency. *Twin Research and Human Genetics*, 9, 367–376. doi:DOI 10.1375/twin.9.3.367 [PubMed: 16790147]
- Rhea SA, Gross AA, Haberstick BC, & Corley RP (2013). Colorado Twin Registry: An update. *Twin Research and Human Genetics*, 16, 351–357. doi:10.1017/thg.2012.93 [PubMed: 23092589]

- Roberts RJ, Hager LD, & Heron C (1994). Prefrontal cognitive processes: Working-memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, 123, 374–393. doi:10.1037//0096-3445.123.4.374
- Rogers RD, & Monsell S (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207–231. doi:10.1037/0096-3445.124.2.207
- Satorra A, & Bentler PM (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507–514. doi:10.1007/Bf02296192
- Schmiedek F, Hildebrandt A, Lövdén M, Wilhelm O, & Lindenberger U (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1089. doi:10.1037/a0015730
- Schoenborn CA, & Heyman KM (2009). Health characteristics of adults aged 55 years and over: United States, 2004–2007. *National Health Statistics Report*, 16, 1–31.
- Shao Z, Janse E, Visser K, & Meyer AS (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, 5, 772. doi:10.3389/fpsyg.2014.00772 [PubMed: 25101034]
- Snyder HR (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin*, 139, 81–132. doi:10.1037/a0028727 [PubMed: 22642228]
- Stroop JR (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. doi:10.1037/0096-3445.121.1.15
- Troyer AK, & Moscovitch M (2006). Cognitive processes of verbal fluency tasks In Poreh AM (Ed.), *The quantified process approach to neuropsychological assessment* (pp. 143–160). New York, NY: Taylor & Francis.
- Tsuang MT, Bar JL, Harley RM, & Lyons MJ (2001). The Harvard Twin Study of Substance Abuse: What we have learned. *Harvard Review of Psychiatry*, 9, 267–279. doi:10.1093/hrp/9.6.267 [PubMed: 11600486]
- Unsworth N, Spillers GJ, & Brewer GA (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *The Quarterly Journal of Experimental Psychology*, 64, 447–466. doi:10.1080/17470218.2010.505292 [PubMed: 20839136]
- Vasilopoulos T, Franz CE, Panizzon MS, Xian H, Grant MD, Lyons MJ, ... Kremen WS (2012). Genetic architecture of the Delis-Kaplan Executive Function System Trail Making Test: Evidence for distinct genetic influences on executive function. *Neuropsychology*, 26, 238–250. doi:10.1037/a0026768 [PubMed: 22201299]
- Wechsler D (1997). *Wechsler adult intelligence scale (3rd ed.)*. San Antonio, TX: Psychological Corporation.
- Wechsler D (1997). *WMS-III: Wechsler memory scale administration and scoring manual*. San Antonio, TX: Psychological Corporation.
- Wechsler D (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: The Psychological Corporation.
- Whiteside DM, Kealey T, Semla M, Luu H, Rice L, Basso MR, & Roper B (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult*, 23, 29–34. doi:10.1080/23279095.2015.1004574 [PubMed: 26111011]
- Wilcox RR, & Keselman HJ (2003). Modern robust data analysis methods: measures of central tendency. *Psychological Methods*, 8, 254–274. doi:10.1037/1082-989X.8.3.254 [PubMed: 14596490]
- Yntema DB (1963). Keeping track of several things at once. *Human Factors*, 5, 7–17. doi:10.1177/001872086300500102 [PubMed: 14002586]

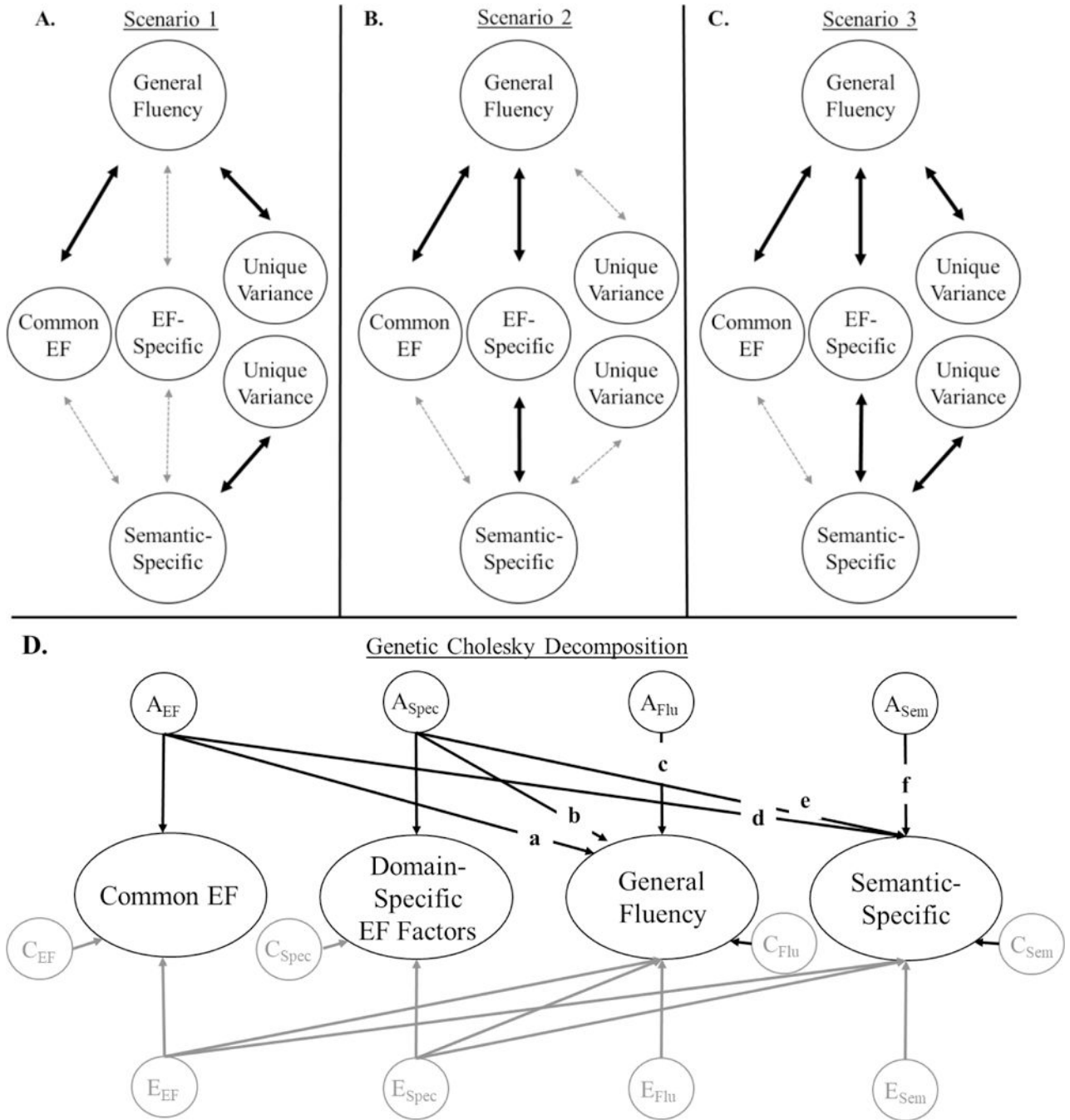
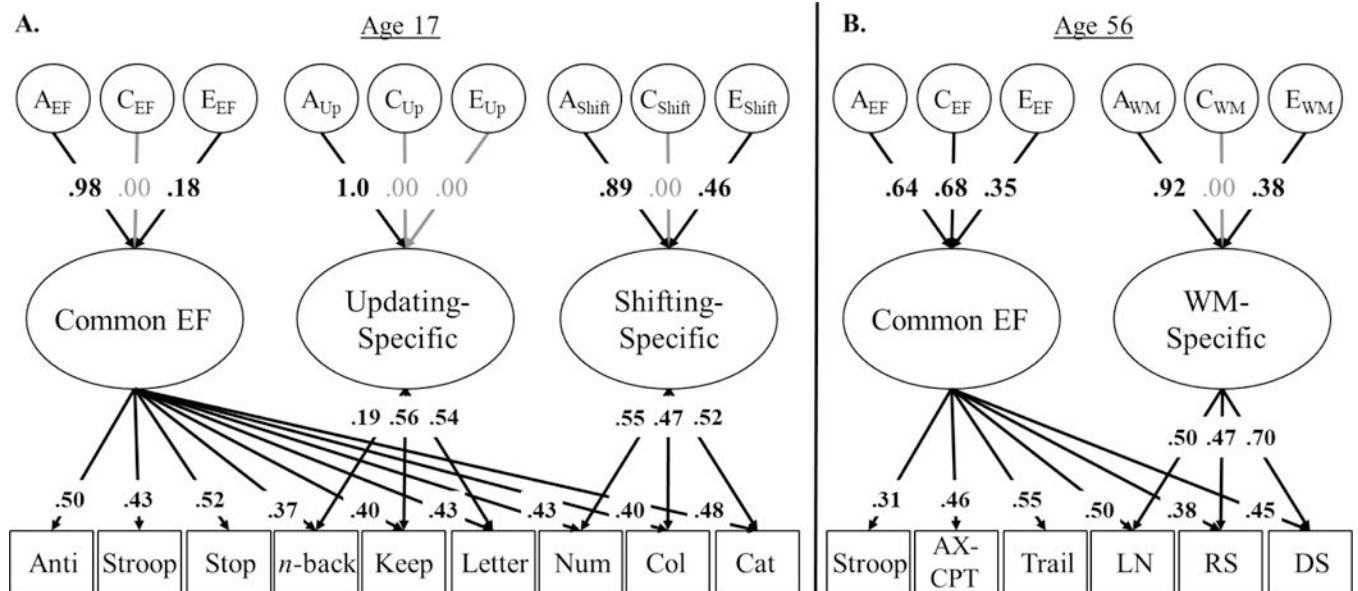


Figure 1: The three hypothesized scenarios of the study (Panels A-C) with Panel D displaying a simplified version of the Cholesky decomposition implemented in genetic analyses. Domain-specific EF factors refer to Updating-Specific and/or Shifting-Specific factors in adolescents and a WM-Specific factor in middle-aged adults. In the first scenario, we expected General Fluency to be associated with Common EF, but also have unique variance (paths a and c, but not path b in Panel D). In the second scenario, we expected General Fluency to be associated with Common EF and one or more domain-specific factors, with no

residual variance (paths a and b, but not c). In the third scenario, we expected General Fluency to be associated with Common EF and one or more EF-specific factors, but also have unique variance (all paths significant). In all scenarios, we expected no association between Common EF and Semantic-Specific factors (path $d = 0$). Panel D depicts the predictions for genetic influences (e.g., A_{EF}), which we expected to account for most of the associations, but we also included shared (C) and nonshared (E) environmental influences, shown in gray (shared environmental covariances were not estimated because they were absent on EF and/or fluency factors). In all models, there are no Cholesky paths between Common EF and domain-specific factors (or between General Fluency and Semantic-Specific fluency factors) because they are uncorrelated by definition. We predicted the same results across samples.

**Figure 2:**

Genetic models of executive function in adolescents and middle-aged adults. The ACE factors represent the genetic (A), shared environmental (C), and nonshared environmental influences (E) on each latent variable. Ellipses indicate latent variables and rectangles indicate measured variables. Percent variance explained can be computed by squaring factor loadings. Residual ACEs on individual measures are not displayed, but were similar to the estimates presented in earlier work (Friedman et al., 2016; Gustavson, Panizzon, Franz, et al., 2018). Both models fit well, $\chi^2(321)=389.00$, $p<.001$, RMSEA=.032, CFI=.956 for Figure 2a, $\chi^2(143)=117.53$, $p=.941$, RMSEA=.000, CFI=1.00 for Figure 2b. Significant factor loadings are displayed in bold and with black lines ($p < .05$).

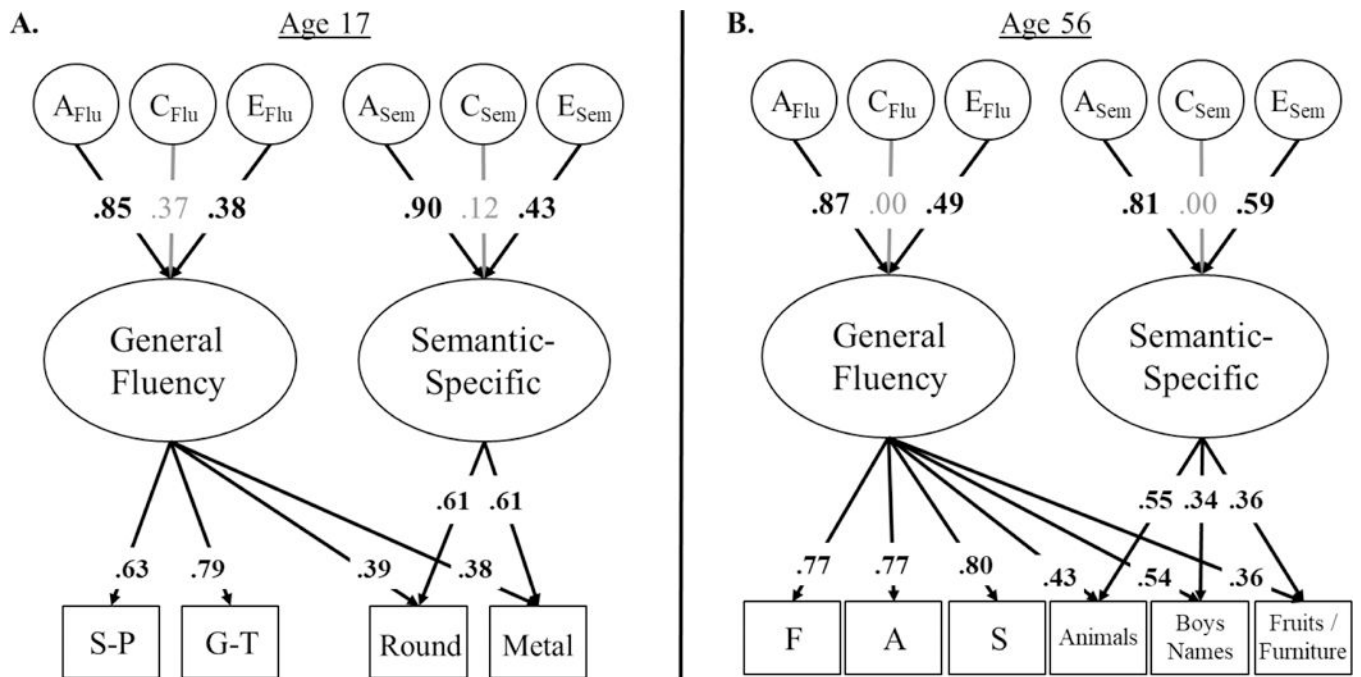


Figure 3:

Genetic models of verbal fluency in adolescents and middle-aged adults. The ACE factors represent the genetic (A), shared environmental (C), and nonshared environmental influences (E) on each latent variable. Ellipses indicate latent variables and rectangles indicate measured variables. Percent variance explained can be computed by squaring factor loadings. Residual ACEs on individual measures are not displayed. Both models fit well, $\chi^2(63)=60.92$, $p=.551$, RMSEA=.000, CFI=1.00 for Figure 3a, $\chi^2(143)=179.54$, $p=.021$, RMSEA=.026, CFI=.986 for Figure 3b. Significant factor loadings are displayed in bold and with black lines ($p < .05$).

Table 1

Sample Characteristics

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Range
<i>Age 17 (CLTS)</i>				
Age (fluency assessment)	811	16.58	0.79	16.0, 20.0
Age (EF assessment)	786	17.25	0.64	16.51, 20.08
Gender (% female)	813	52	-	-
Intelligence	812	102.17	11.41	70, 142
<i>Age 56 (VETSA)</i>				
Age	1290	55.9	2.44	51.1, 60.7
Education (years)	1290	13.85	2.1	5, 20
General Cognitive Ability (age 20)	1274	0.34	0.69	-1.29, 2.32
General Cognitive Ability (age 56)	1288	0.42	0.64	-1.76, 2.05

Note: In the CLTS, intelligence was measured with the WAIS-III. In VETSA, general cognitive ability was assessed with all four subscales of the Armed Forces Qualifications Test, which were transformed and percentile scored based on military norms. The means correspond to an IQ of approximately 104 to 105 on the WAIS (Lyons et al., 2009; Orme, Brehm, & Ree, 2001). Thus, IQ is representative of the general population in both samples.

Table 2

Descriptive Statistics

Task	<i>N</i>	<i>M</i>	<i>SD</i>	Range	Skewness	Kurtosis
<i>Age 17 (CLTS)</i>						
Stroop	759	213.60	90.11	0, 488	0.58	0.19
Antisaccade ^a	779	1.04	0.20	.47, 1.57	-0.12	-0.26
Stop Signal	741	281.92	62.56	151, 489	1.13	1.51
Keep Track ^a	774	0.94	0.18	.38, 1.49	0.30	0.56
Letter Memory ^a	785	1.09	0.25	.38, 1.57	0.29	-0.20
<i>n</i> -back ^a	777	1.17	0.17	.65, 1.57	-0.92	1.65
Number-Letter	776	330.80	183.18	-14, 923	1.04	1.12
Color-Shape	768	331.18	189.42	-196, 916	0.76	0.75
Category Switch	766	333.27	181.38	-34, 899	0.98	0.92
Phonemic S-P	811	8.39	3.90	0, 21	0.33	-0.15
Phonemic G-T	811	5.71	2.31	1, 14	0.50	-0.04
Semantic Round	811	13.50	4.90	2, 36	0.61	0.59
Semantic Metal	811	14.67	5.98	2, 35	0.35	-0.24
Vocabulary Part 1	807	24.39	8.43	1, 49	0.47	0.10
Vocabulary Part 2	808	5.83	2.95	0, 24	1.20	3.86
Performance IQ ^b	813	101.01	11.17	72, 136	0.20	-0.11
<i>Age 56 (VETSA)</i>						
Stroop ^c	1255	35.95	8.33	6, 69	0.05	0.30
AX-CPT ^d	1190	0.99	0.37	0, 1.39	-1.35	1.06
Trail Making Test ^c	1271	89.15	35.05	31.34, 240.00	1.68	3.87
Letter-Number	1280	10.13	2.36	0, 20	0.22	1.02
Reading Span	1248	34.14	5.33	15, 45	-0.54	0.05
Digit Span	1276	17.11	3.90	8, 30	0.29	-0.31
Phonemic F	1277	12.28	4.09	1, 29	0.29	-0.02
Phonemic A	1277	11.15	3.90	1, 29	0.41	0.34
Phonemic S	1277	13.48	4.32	1, 31	0.25	0.08
Animals	1275	19.20	4.43	6, 39	0.26	0.26
Boys' Names	1276	19.09	4.48	6, 40	0.34	0.58
Fruits/Furniture	1277	12.75	2.55	4, 22	-0.01	0.30
WASI - Vocabulary	1284	54.77	8.25	21, 72	-0.62	0.67
AFQT - Vocabulary ^d	1203	83.05	17.82	1, 99	-1.72	2.73
WASI - Matrix Reasoning	1288	54.62	8.97	21, 72	-1.20	1.23
AFQT - Nonverbal ^d	1288	58.91	20.86	4, 93	-0.27	-0.66

Note: In all analyses, dependent measures were standardized residual scores after adjusting for age (and sex for the adolescent sample only), but the unadjusted scores are presented here. Reaction time-based measures are also displayed before multiplying by negative 1 in all analyses (so higher scores always indicates better performance).

^a indicates the dependent measure was arcsine-transformed.

^b indicates the dependent measure was a scaled score.

^c indicates performance on the EF condition before being adjusted for performance on baseline trials (color and word trials for Stroop, number only or letter only for Trail Making Test).

^d indicates the dependent measure was a percentile score. AX-CPT = AX – Continuous Performance Test; WASI = Wechsler Abbreviated Scale of Intelligence; AFQT = Armed Forced Qualification Test.

Table 3

Latent Variables Correlations between Verbal Fluency, Executive Functions, Vocabulary, and Nonverbal Cognitive Ability

	General Fluency	Semantic-Specific	Common EF	Updating-Specific	Shifting-Specific	Vocabulary
<i>Age 17</i>						
General Fluency	1					
Semantic-Specific	-	1				
Common EF	0.43	-0.03	1			
	[.28, .58]	[-.19, .14]				
Updating-Specific	0.48	0.28	-	1		
	[.33, .64]	[.12, .44]				
Shifting-Specific	-0.18	0.39	-	-	1	
	[-.33, -.03]	[.21, .57]				
Vocabulary	0.57	0.32	0.32	0.54	0.05	1
	[.47, .66]	[.23, .40]	[.20, .44]	[.40, .67]	[-.07, .16]	
Nonverbal Cognitive Ability	0.40	0.30	0.48	0.27	-0.14	0.39
	[.32, .48]	[.20, .40]	[.38, .57]	[.13, .39]	[-.26, -.02]	[.31, .46]
	General Fluency	Semantic-Specific	Common EF	WM-Specific		Vocabulary
<i>Age 56</i>						
General Fluency	1					
Semantic-Specific	-	1				
Common EF	0.41	0.18	1			
	[.31, .50]	[.05, .32]				
WM-Specific	0.33	-0.04	-	1		
	[.24, .42]	[-.17, .09]				
Vocabulary	0.50	0.22	0.59	0.34	-	
	[.44, .56]	[.14, .31]	[.49, .68]	[.24, .43]		
Nonverbal Cognitive Ability	0.35	0.26	0.72	0.11	-	0.68
	[.27, .42]	[.15, .36]	[.62, .82]	[.00, .22]		[.61, .75]

Note: Significant correlations are displayed in bold ($p < .05$). All constructs were measured with latent variables except nonverbal cognitive ability in adolescents (i.e., WAIS performance IQ). 95% confidence intervals are presented in brackets. Models fit the data well at both waves, $\chi^2(80)=188.03$, $p < .001$, RMSEA=.041 for adolescents, $\chi^2(82)=134.64$, $p < .001$, RMSEA=.022 for middle-aged adults.

Table 4

Regression Analyses Involving Executive Functions, Vocabulary, and Nonverbal Cognitive Ability

	Common EF	Updating-/ WM-Specific	Shifting- Specific	Vocabulary	Nonverbal Cognitive Ability	R ²
<i>Age 17</i>						
General Fluency	0.32 [.15, .49]	0.32 [.14, .50]	-0.19 [-.34, -.04]	0.29 [.14, .44]	0.02 [-.09, .14]	0.50
Semantic-Specific	-0.21 [-.43, .00]	0.13 [-.09, .35]	0.39 [.17, .60]	0.25 [.08, .43]	0.40 [.24, .57]	0.43
<i>Age 56</i>						
General Fluency	0.29 [.12, .46]	0.23 [.13, .31]	-	0.34 [.21, .46]	-0.12 [-.30, .06]	0.33
Semantic-Specific	-0.01 [-.27, .24]	-0.08 [-.22, .06]	-	0.19 [.01, .38]	0.19 [-.05, .42]	0.11

Note: Standardized regression coefficients and their 95% confidence intervals are reported. In these regression models, fluency factors were included as dependent variables with correlations among all independent variables (EF factors, vocabulary, nonverbal cognitive ability), except correlations among EF factors were fixed to zero by definition (e.g., Common EF and WM-Specific at age 56). Correlations among independent variables were identical to those from the correlational models (Table 3). All constructs were measured with latent variables except nonverbal cognitive ability in adolescents (i.e., WAIS performance IQ). Significant regression coefficients are displayed in bold ($p < .05$).

^a indicates Updating-Specific at age 17 but Working Memory-Specific at age 56.

Table 5

Genetic and Environmental Correlations Between Verbal Fluency and Executive Functions

Fluency Measure	Common EF		Updating-/WM-Specific ^a		Shifting-Specific	
	r_g	r_e	r_g	r_e	r_g	r_e
<i>Age 17</i>						
General Fluency	0.37 [.22, .54]	0.84 [.27, .99]	0.54 [.39, .74]	-	-0.11 [-.30, .07]	-0.54 [-.86, -.11]
Semantic-Specific	0.03 [-.21, .24]	0.14 [-.39, .76]	0.38 [.11, .64]	-	0.30 [.06, .57]	0.58 [.15, 1.0]
<i>Age 56</i>						
General Fluency	0.71 [.47, .95]	0.20 [-.17, .58]	0.40 [.26, .58]	0.10 [-.29, .46]	-	-
Semantic-Specific	0.53 [.11, .87]	0.27 [-.27, .93]	0.02 [-.26, .26]	0.17 [-.41, .72]	-	-

Note: Genetic (r_g) and environmental (r_e) correlations between executive functions and verbal fluency (and their 95% confidence intervals) are displayed in adolescents (top) and middle-aged adults (bottom). The total variance explained by genetic (a^2) and nonshared environmental influences (e^2) were nearly identical to those displayed in Figures 2 and 3. Shared environmental correlations were not estimated because, in all cases, one or both variables had no evidence for shared environmental influences. Nonshared environmental correlations between the Updating-Specific factor and fluency factors (age 17) were not estimated because the environmental estimate on the Updating-Specific factor was also zero (see Figure 2a). Genetic and environmental correlations were not estimated directly but computed from the output of the Cholesky decomposition. Factor loadings and residual ACEs on individual tasks are also not displayed, but were similar to those in Figures 2 and 3. Also not displayed, there were implied correlations between General Fluency and Semantic-Fluency through their overlap with EFs ($r_g=.18$, $r_e=-.20$ at age 17, $r_g=.39$, $r_e=.07$ at age 56). Both models fit well, $\chi^2(663)=826.24$, $p<.001$, RMSEA=.034, CFI=.938 for age 17, $\chi^2(565)=637.92$, $p=.018$, RMSEA=.019, CFI=.985 for age 56. Significant correlations are displayed in bold ($p<.05$).

^a indicates Updating-Specific at age 17 but Working Memory-Specific at age 56.