



Published in final edited form as:

J Exp Psychol Gen. 2022 August ; 151(8): 1745–1761. doi:10.1037/xge0001168.

Genetic Associations Between Executive Functions and Intelligence: A Combined Twin and Adoption Study

Daniel E. Gustavson¹, Chandra A. Reynolds², Robin P. Corley³, Sally J. Wadsworth³, John K. Hewitt³, Naomi P. Friedman^{3,4}

¹Department of Medicine, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN

²Department of Psychology, University of California Riverside, Riverside, CA

³Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO

⁴Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO

Abstract

Much debate has concerned the separability of executive function abilities and intelligence, with some evidence that the two constructs are genetically indistinguishable in children and adolescents but phenotypically and genetically distinct in older adolescents and adults. The current study leveraged data from twin and adoption studies to examine executive function's genetic structure in adulthood ($M=33.15$ years, $SD=4.96$) and its overlap with intelligence. 1238 individuals (170 MZ twin pairs, 154 DZ twin pairs, 95 biological sibling pairs, 80 adoptive sibling pairs, and 240 unpaired individuals) completed 6 executive function tasks as well as the Weschler Adult Intelligence Scale-III as part of the Colorado Adoption/Twin Study of Lifespan behavioral development and cognitive aging (CATSLife). Results replicated the unity/diversity model of executive function that distinguishes general executive function abilities (Common EF) from abilities specific to working memory updating (Updating-specific) and mental set shifting (Shifting-specific). In the final model, broad-sense heritability was high for Common EF ($h^2=.72$), Updating-specific ($h^2=1.0$), and Shifting-specific ($h^2=.60$) factors, as well as for full-scale intelligence ($h^2=.74$). Intelligence was phenotypically and genetically correlated with Common EF ($r=.49$, broad-sense $r_g=.44$) and Updating-specific ($r=.60$, $r_g=.69$) abilities. This study represents the first executive function study to apply the adoption design. Leveraging the combined twin and adoptive design allowed us to estimate both additive and nonadditive genetic effects underlying these associations. These findings highlight the commonality and separability of executive function and intelligence. Common EF abilities are distinct from intelligence in adulthood, with intelligence also strongly associated with Updating-specific abilities.

Correspondence concerning this article should be addressed to Daniel Gustavson, Department of Medicine, Vanderbilt University Medical Center, 2525 West End Ave., Suite 700, Nashville, TN, 37232. daniel.e.gustavson@vumc.org.

We have no known conflicts of interest to disclose.

None of the data or ideas described in this manuscript have been disseminated prior to appearing in this manuscript. As the data are from a large longitudinal study, subsets of the data have been included in prior articles as described in the Sample Overlap with Prior Work section.

All stimulus materials and data are available upon request. Additionally, input and output scripts for all phenotypic invariance analyses (in Mplus) and genetic analyses (in R Markdown) are available at the following webpage: https://osf.io/rk6yv/?view_only=accb098c2842415e9f47769721888cf0

Keywords

heritability; executive control; cognitive control; twin study; adoption study

Executive functions (EFs) are a set of domain-general cognitive control processes that enable goal-directed behavior (Miyake & Friedman, 2012; Miyake et al., 2000). EFs are typically assessed with measures of distinct yet correlated abilities, including prepotent response inhibition, working memory updating, and mental set shifting (Miyake et al., 2000). A Common EF factor captures the general covariation across multiple EF components. The constructs of intelligence or general cognitive ability (g) were developed earlier to account for the variance across a much broader range of cognitive tasks (Neisser et al., 1996; Spearman, 1904). For many years, researchers have debated whether EFs and intelligence represent truly distinct abilities, or if the cognitive control abilities captured by EFs are encapsulated by the general abilities reflected in measures of intelligence or g (Kyllonen & Christal, 1990; Salthouse, 2005). Genetic studies can inform this debate by quantifying construct similarity at different levels of analysis: genetic and environmental.

Here, we address the commonality and separability of EFs and intelligence at the phenotypic and genetic/environmental levels by leveraging latent variable analyses of data from a combined twin and adoption study of adults. We test the hypothesis that Common EF abilities are only moderately correlated with intelligence measures in adulthood, with additional correlations between intelligence and variance specific to working memory updating ability. Combining twin and adoption samples provides a unique opportunity to simultaneously quantify two sources of genetic influences (additive, dominance) and two sources of environmental influences (shared and nonshared within families), as well as how these influences on EFs and intelligence covary (c.f., Matteson, McGue, & Iacono, 2013; Plomin & DeFries, 1985). Specifically, in addition to the additive genetic and nonshared environmental influences that are consistently observed for both EFs and intelligence, we evaluate the hypothesis that intelligence has some small shared environmental influences whereas EFs are instead influenced by dominance (nonadditive) genetic influences.

Unity and Diversity Framework of Executive Function

The term EF is used to capture a wide range of abilities that contribute to the regulation of thoughts and behavior. The most frequently studied EFs in latent variable models are stopping automatic or dominant responses (response Inhibition), Updating working memory, and Shifting between mental sets (Miyake & Friedman, 2012; Miyake et al., 2000). Miyake et al. (2000) demonstrated that latent factors for Inhibition, Updating, and Shifting share considerable variance, referred to as the “unity” of EFs. However, they also represent separable processes ($r_s = .42$ to $.63$, significantly less than 1.0), demonstrating the “diversity” of EFs. The more recent unity and diversity framework of EFs is a reparameterization of the original correlated-factors approach in which variance across EF tasks is decomposed into three orthogonal latent factors: Common EF, Updating-specific ability, and Shifting-specific ability (Friedman et al., 2008). As yet, there is no evidence for inhibition-specific variance,

suggesting that all the EF-relevant individual differences in response inhibition are captured by Common EF.

Common EF is thought to reflect the ability to create, maintain, and use goals to bias ongoing cognitive processes, as these general goal-management processes are needed across multiple situations. Updating-specific ability is thought to reflect the selective gating of information in to and out of working memory, as well as memory retrieval. Shifting-specific ability is thought to reflect the ability to flexibly replace goals that are no longer necessary (Friedman & Miyake, 2017; Miyake & Friedman, 2012). Researchers have demonstrated that Common EF is associated with a wide range of important cognitive, personality, and mental health traits (Friedman, du Pont, Corley, & Hewitt, 2018; Gustavson et al., 2019; Gustavson, Miyake, Hewitt, & Friedman, 2015; Ito et al., 2015). In contrast, Updating-specific and Shifting-specific abilities are typically not strongly associated with these other traits, supporting the idea that the Common EF factor captures the aspects of EF most relevant to everyday cognitive control, self-restraint, and goal management (Gustavson et al., 2015).

Executive Functions and Intelligence

Early views of EFs and intelligence highlighted that EF-related traits such as inhibition-related processes are important components of intelligence (Dempster, 1991) and demonstrated especially strong correlations between working memory capacity and intelligence (Carpenter, Just, & Shell, 1990; Colom, Rebollo, Palacios, & Juan-Espinosa, 2004; Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen & Christal, 1990). Others, using a neuropsychological approach, highlighted the fact that patients with frontal lobe damage (often associated with EF deficits) also showed impaired fluid intelligence, but not crystallized intelligence (Duncan, Burgess, & Emslie, 1995; Duncan et al., 1996). Since Miyake et al.'s (2000) study, more individual differences work has focused on the overlap between EFs and intelligence using individual measures or latent variables capturing Inhibition, Shifting, and Updating. This work has confirmed that intelligence measures are correlated with each of these aspects of EF (Friedman et al., 2006; Salthouse, Atkinson, & Berish, 2003; Unsworth et al., 2009).

Genetic studies have also been useful in understanding individual differences in EFs and their relations to intelligence. For example, in late adolescence, genetic influences appear to explain the vast majority of the variance in Common EF (98%), Updating-specific (99%), and Shifting-specific (79%) latent variables, with considerable genetic correlations between Common EF and full scale intelligence (additive genetic correlation [r_a]=.57) and between Updating-specific ability and intelligence (r_a =.56) (Friedman et al., 2008). The Friedman et al. (2008) study was important in demonstrating that the genetic influences on intelligence are not identical to those on Common EF. Rather, intelligence was also associated with Updating-specific variance, in line with earlier proposals that intelligence appears tightly linked with working memory capacity (Kyllonen & Christal, 1990) but not necessarily other EF abilities. In other words, working memory capacity and updating ability may be most strongly associated with intelligence because it is a combination of Common EF and Updating-specific processes: Common EF is necessary to establish a strong task

representation and focus attention, and Updating-specific ability is necessary to support working memory operations such as addition of new information or removal of irrelevant information. Both processes are strongly – yet independently – associated with intelligence.

Similar results were recently observed in a middle-aged sample (Gustavson et al., 2018b). A latent Common EF factor based on seven EF tasks was phenotypically correlated with a measure of general cognitive ability ($r=.54$, $r_a=.59$), but so was a latent factor capturing variance specific to working memory ($r=.09$, $r_a=.24$). The smaller correlation with working memory-specific processes in this study compared to the Friedman et al. (2008) study was likely due to the reduced demands on updating in neuropsychological working memory tasks (Gustavson et al., 2018b), consistent with earlier research demonstrating stronger associations with intelligence as demands on working memory updating (vs. working memory span) increase (Engle et al., 1999). Nevertheless, this work demonstrates independent associations between intelligence, Common EF, and working memory abilities. Across both studies, the strong genetic correlations but weak environmental correlations indicate that EFs and intelligence share considerable genetic influences but are differentiated by environmental influences.

Shifting-specific ability has only been examined in relation to intelligence in one earlier investigation (Friedman et al., 2008). In contrast to the results for Updating-specific, Shifting-specific was weakly negatively associated with concurrent intelligence in this study at the genetic level ($r_a = -.20$). These findings were intriguing in that they suggest that the flexibility thought to be captured by Shifting-specific factor — the ease of transitioning to new task-set representations (Goschke, 2000; Miyake & Friedman, 2012) — could be at odds with Common EF processes or intelligence tests that require individuals to strongly engage a task set and maintain this task representation. However, studies have not yet replicated this result, and it is possible that Shifting-specific ability and intelligence are simply unrelated.

A twin study of children and adolescents (mean age 11 years) demonstrated strong phenotypic and genetic overlap between a latent Common EF factor and full-scale intelligence ($r=.71$; $r_a=.92$) (Engelhardt et al., 2016). Although these findings suggest that Common EF and intelligence are nearly indistinguishable at the genetic level, this hierarchical model did not distinguish Common EF from Updating-specific variance. In fact, their Common EF factor had higher factor loadings on their Updating (.92) and Working Memory span (.80) subfactors compared to the Inhibition (.52) and task Switching (.73) factors). Another study of similarly aged children (mean age 10 years), including twins, from the Adolescent Brain and Cognitive Development study also found a high genetic correlation between Common EF and intelligence ($r=.64$; $r_a=.86$) which did not significantly differ from 1.0 (Freis et al., 2021). However, this study also found that intelligence was correlated with the small amount of variance specific to working memory updating tasks in that sample ($r=.81$; $r_a=.36$), over and above its relation to Common EF.

The factor structure of EFs appears to be unitary in children before about age 10–12 (Brydges, Fox, Reid, & Anderson, 2014; Hartung et al., 2020), suggesting that potentially independent genetic associations between intelligence and Common EF versus Updating-

specific ability may only emerge in late adolescence and adulthood. Such findings are consistent with another phenotypic study demonstrating strong associations of a unitary EF factor with fluid and crystallized intelligence in children (Brydges, Reid, Fox, & Anderson, 2012).

A key way that genetic studies can be utilized to understand the difference between EFs and intelligence is by examination of shared environmental versus dominance genetic influences. Although shared environmental influences are small for intelligence, they reliably explain 10–25% of the variance (Engelhardt et al., 2016; Friedman et al., 2008; Lyons et al., 2009; Rowe, Vesterdal, & Rodgers, 1998). In contrast, studies of EF have found little evidence for shared environmental influences on Common EF or EF-specific factors in adolescents or young adults (Engelhardt et al., 2016; Friedman et al., 2016; Friedman et al., 2008). Instead, twin correlations from these studies suggest there may be both additive and dominance genetic variance on EF abilities (i.e., the within-trait correlations among monozygotic [MZ] twin pairs are greater than 2 times the correlation among dizygotic [DZ] twin pairs). An exception is that in one sample of middle-aged adults, shared environmental influences on Common EF were observed and correlated with those for general cognitive ability (Gustavson et al., 2018b). However, twin correlations for individual EF tasks in this sample showed some evidence for dominance genetic influences at both waves (mean age 56 & 62 years; Gustavson et al., 2018a).

Unfortunately, with a classic twin design involving only MZ and DZ twin pairs, it is impossible to estimate both shared environmental influences and dominance genetic influences in the same model. Researchers typically chose ACE models (i.e., Additive genetic, shared/Common environment, and nonshared Environmental influences) over ADE models (with Dominance genetic influences instead of common environmental influences) because dominance genetic influences are inconsistently suggested by twin correlations and their detection is underpowered even in large samples (Martin, Eaves, Kearsy, & Davies, 1978). However, twin correlations are imperfect indicators of these influences, because if both dominance genetic variance and shared environmental variance influence a trait and only one is estimated, these two influences actually cancel one another out in the estimates of C only or D only. This cancelling out occurs because shared environmental influences tend to make DZ twin correlations larger, whereas dominance genetic influences tend to make DZ twin correlations smaller, in comparison to MZ twin correlations.

Estimating a full ADCE model requires an extended family design that includes another group with a different degree of relatedness, with a group of unrelated adopted siblings providing the greatest increase in power (Plomin & DeFries, 1985). Indeed, even though twin correlations typically support shared environmental influences on intelligence, when an extended family design is used there is evidence for weak dominance genetic variance as well (Keller et al., 2013). Because EFs and intelligence measures are both highly heritable, but may differ in the magnitude of their dominance genetic and shared environmental influences, being able to estimate all four variance components in the context of a single model will shed more light on the commonality and separability of these constructs.

As a final point, genetic studies of EFs in adulthood are almost entirely absent. Quantifying the heritability of EF factors and their overlap with intelligence in this adult period (late 20s to late 40s), when frontal lobes have fully developed but cognitive decline may have not yet begun, will provide an important data point in the development of this relationship. It will also shed light on whether shared environmental influences for Common EF begin to reveal themselves earlier than middle age (Gustavson et al., 2018a).

Hypotheses of the Current Study

We examined phenotypic and genetic/environmental associations of Common EF, Updating-specific, and Shifting-specific factors with intelligence in combined twin and adoption studies from the Colorado Adoption/Twin Study of Lifespan behavioral development and cognitive aging (CATSLife) project (Wadsworth et al., 2019). We predicted that intelligence would be associated with both Common EF and Updating-specific abilities, and that these associations would be primarily driven by genetic correlations. Given the high heritability estimates from earlier twin samples, we expected EF factors would be explained by both additive and dominance genetic variance, especially Updating-specific ability, which was almost entirely explained by genetic influences (Friedman et al., 2016). An open question concerned whether there would be unique genetic/environmental variance on intelligence after accounting for those genetic and environmental influences on Common EF and Updating-specific ability.

Method

Participants

Analyses were based on 1238 individuals ($M=33.15$ years, $SD=4.96$) who completed measures of EF and intelligence as part of the ongoing Colorado Longitudinal Twin Study (LTS) EF-MRI (Corley et al., 2019) and CATSLife projects (Wadsworth et al., 2019). This total included 700 individuals from same-sex twin pairs from the LTS ($M=29.33$ years, $SD=1.27$) and 538 individuals from the Colorado Adoption Project (CAP; $M=38.12$ years, $SD=3.27$), and included EF data collected for LTS EF-MRI study or CATSLife between 2014–2020 and intelligence data collected in CATSLife between 2015–2020. These studies were approved by the institutional review board at the University of Colorado Boulder.

The LTS subsample was originally 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 (91.6%), Hawaiian, Pacific islander, American Indian, or Alaskan (1.6%), more than one race (5.4%), or unknown or unreported race (1.4%). Hispanic individuals comprised 9.7% of the sample. For more information on the sample characteristics, see Corley et al. (2019). Genetic analyses included 170 full MZ twin pairs, 154 full DZ twin pairs, and 52 unpaired twins.

The CAP subsample was recruited beginning in 1975 through two Denver adoption agencies (Plomin & DeFries, 1983). Parents were recruited with a one-to-one ratio of adoptive and nonadoptive parents. Initially, the first younger sibling in the family was also enrolled, but later studies expanded to include other siblings. Participants identified as White (92.2%),

Asian (5.0%), American Indian or Alaskan (1.3%), Black or African American (0.6%), or more than one race (0.9%). Hispanic individuals comprised 0.9% of the sample. For preliminary phenotypic analyses and data harmonization, all siblings were included. For genetic/environmental models, we chose siblings to pair with each proband based on the order of enrollment (i.e., first male or female sibling). If they had both a male and female sibling, we chose the sibling with the most data. If both had complete data, we chose the same-sex sibling. In these analyses, we had 95 biological sibling pairs, 80 adoptive sibling pairs, and 188 unpaired siblings (including 21 individuals excluded in genetic analyses for being the 3rd sibling in a family or different type of relationship [i.e., half-sibling]).

Procedure

Participants completed the study in one of two ways: All individuals from the CAP and a subset of LTS completed EF and intelligence measures as part of the first wave of CATSLife (collected from 2015 to 2020). The remaining LTS participants ($n=485$) completed 3 of the EF tasks (antisaccade, keep track, number-letter switching) in an MRI scanner and the remaining 3 EF tasks outside the scanner as part of a separate assessment (LTS EF-MRI) $M=0.17$ years ($SD=.77$) before the primary CATSLife assessment (Corley et al., 2019). These individuals did not redo the EF tasks during CATSLife because the assessments were close in time. Some individuals completed EF tasks as part of both assessments ($n=175$) when EF-MRI and CATSLife were completed more than 1 years apart, so scores on the MRI and non-scanner versions of these 3 EFs tasks were harmonized based on these individuals (see Data Analysis). For these 175 individuals, we analyzed data based on their CATSLife assessment, which was completed $M=1.75$ years ($SD=.89$) after the EF-MRI assessment. The exception was the letter memory task, which showed evidence for practice effects, $t(170)=8.50$, $p<.001$, so we kept the first assessment (typically, EF-MRI). Intelligence was assessed during the CATSLife assessment for all subjects.

Measures

The measures reported here represent a subset of those available from the EF-MRI and CATSLife studies, which are both large-scale studies of cognition and health that build on decades of longitudinal assessments. In this section, we describe only those measures related to EF and intelligence (and only those collected at the most recent assessment for CATSLife), as all other measures were outside the scope of this investigation.

Executive function.—Six EF tasks, two per latent construct (Inhibition, Updating, and Shifting), were selected from the original set of nine tasks used in prior studies with the LTS sample (see Friedman et al., 2016). These six tasks were selected because they had the highest factor loadings on their respective factors. The scanner versions of the tasks were similar to the laboratory versions, in that they were designed to elicit individual differences in performance, but also to enable MRI analysis (i.e., including jittered timings and comparison conditions).

Inhibition.: The two inhibition tasks required participants to stop a dominant or prepotent response. The color-word *Stroop task* (Stroop, 1935) required participants to avoid the dominant tendency to read words, and instead name the colors in which the words were

printed (on a black background). This task, which was completed in the laboratory for all participants, was the same one used at the age 23 LTS assessment (Friedman et al., 2016). On each trial, a 750 ms blank period was followed by a 250 ms white fixation cross, then a colored target, which remained on the screen until the participant reported the color. RTs were collected by using voice key, and stimuli disappeared as soon as the voice key detected the response. The task began with 10 practice and 42 actual neutral trials (colored strings of 3–5 asterisks), followed by a block of 10 practice and 42 actual congruent trials (RED, BLUE, and GREEN printed in the congruent color), and ended with two blocks of 42 incongruent trials (color words printed in incongruent colors). Each block included two warm-up trials that were not included in the analyses. The dependent measure was reaction time (RT) interference, obtained by subtracting mean RT for correct asterisks trials from mean RT correct incongruent trials.

In the *antisaccade* task (Roberts, Hager, & Heron, 1994), participants had to avoid a reflexive tendency to saccade to a cue stimulus and instead immediately look to the opposite side of the screen in time to identify a digit (1–9) that appeared briefly before being masked. The non-scanner version was a shortened version of the one used at the age 23 LTS assessment (Friedman et al., 2016). Participants completed 12 practice and 18 actual prosaccade trials to familiarize themselves with the task and reinforce the prepotent response, followed by 12 practice antisaccade trials then two blocks (36 antisaccade trials each) that increased in difficulty. Each trial began with a centered fixation cross that remained on the screen for one of nine durations from 1,500 to 3500 ms (in 250-ms increments). The subsequent cue was a black 1/8 in. square (inner edge 3.375 in. from the center), which appeared on the left or right of the screen with equal probability. The cue remained on the screen for 233 ms in the first antisaccade block and 200 ms in the second antisaccade block and the initial prosaccade blocks. Once it disappeared, a numeric target (a digit 1–9, 26 pt Helvetica font, presented in a 7/16 in. square with its inner edge 3.25 in. from the fixation) appeared for 150 ms before being masked, on the same side as the cue (prosaccade block) or on the opposite side (antisaccade blocks). The mask remained on the screen until the participant verbalized the target number (or guessed), and the experimenter entered the response, initiating the next trial. Each block contained two warm up trials that were not included in the analyses.

The scanner version of the antisaccade task consisted of shorter blocks (20 s each) of antisaccade trials that were intermixed with comparison prosaccade and fixation trials (24 blocks of each; 5 trials per block for the prosaccade and antisaccade blocks). Each block was preceded by an instruction (TOWARD, AWAY, or FIXATION) indicating the trial type. Each trial began with a fixation lasting 1,000–3,000 ms (in 500 ms intervals), and the cue lasted for 233 ms, however, the duration of this cue was changed to 283 ms after the first 276 participants as we noticed low average performance in an interim analysis for another project. Targets were digits 0–9, and masks lasted 1650 ms, during which time the participant was instructed to vocalize the target. The dependent measure was the total accuracy across 72 antisaccade trials in the laboratory version and 60 antisaccade trials in the fMRI version.¹

Shifting: The two shifting tasks required participants to switch between categorization dimensions according to a cue that appeared 350 ms before the stimulus and remained on the screen along with the target stimulus until the participant entered 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 cue symbol (a heart or a cross) that appeared above the word. This task, which was completed in the laboratory for all participants, was the same one used at the age 23 LTS assessment (Friedman et al., 2016). In the *number-letter task* (Rogers & Monsell, 1995), participants used two buttons on a ms-accurate button box to indicate whether number-letter or letter-number 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). The cue was the darkening of the border around the quadrant in which the target would appear. The dependent measures for both tasks were local switch costs based on blocks in which the two tasks were intermixed: The mean RT during trials that required a switch between categorization rules minus the mean RT during trials where the same rule was repeated (64 trials per condition in the behavioral tasks, 96 per condition in the scanner version of the number-letter task).

The category switch task and the laboratory version of the number-letter task consisted of 12 practice then 32 actual task 1 trials (living/nonliving or number), 12 practice then 32 actual task 2 trials (small/big or letter), then 24 practice and 2 blocks of 64 actual mixed task trials (50% required a task switch). There was a 350 ms response to cue interval, and a 200-ms buzz sounded for errors (buzzes were removed for the scanner version). Each block included two (single-task blocks) or four (mixed blocks) warm-up trials that were not included in the analyses. In the scanner version of the number-letter task, the trials were arranged in shorter blocks of 13 trials (12 trials + 1 warm-up trial) each, which were intermixed with 20-s rest blocks. Each block was preceded by an instruction indicating the trial type (TOP, BOTTOM, MIXED, or FIXATION). In mixed blocks, half the trials were switch trials. The task contained 16 mixed blocks, 8 single-task blocks, and 8 fixation blocks.

Updating: The two updating tasks required participants to monitor and continuously manipulate the contents of working memory. The *letter memory task* (Morris & Jones, 1990) was completed in the laboratory for all participants and was the same one used at the age 23 LTS assessment (Friedman et al., 2016). Participants saw a list of letters presented one at a time for 3 seconds. With each letter, participants said aloud the four most recent letters that appeared. The dependent measure was the proportion of sets correctly rehearsed across 12 trials (four trials each of length 9, 11, or 13 letters; 132 total sets). Points were given for rehearsing only the correct letters in the correct serial order.

In the *keep-track task* (Yntema, 1963), participants saw a list of 15 (laboratory) or 16 (scanner) words presented one at a time on the screen at a rate of 2 seconds per word.

¹This was a speeded version of the antisaccade task, for which individual differences arise in the accuracy score. Because performance is typically at ceiling for prosaccade trials, prosaccade trials are not included in the performance score, consistent with the dependent measures used in previous work validating this model (Friedman et al., 2008; Miyake et al., 2000).

The words were drawn from a list of 36 words, 6 words in each of 6 categories (animals, colors, countries, distances, metals, and relatives). They were instructed to remember only the most recent word presented from the target specified categories (displayed on the bottom of the screen during the entire trial) and respond at the end of the trial when a prompt appeared (“???”). In the scanner version, the response period lasted for 10 s during which the participant vocalized answers. In the laboratory version, participants also vocalized their answers, but there was no time limit. Recall trials were intermixed with fixation blocks and comparison READ trials in which participants were instructed to read words without trying to remember them. The dependent measure was the proportion of total words recalled across all 9 trials (36 words in the laboratory version: three trials each of three, four, and five categories; 33 words in the scanner version: three trials of three categories and six trials of four categories).

Intelligence.—Intelligence was assessed with the Wechsler Adult Intelligence Scale, third edition (WAIS-III; Wechsler, 1997). We used the full-scale intelligence quotient (IQ) score in all analyses.

Harmonization and Controlling for Demographic Characteristics

To create harmonized scores for the $n=485$ individuals who completed the antisaccade, keep track, and number-letter tasks in the MRI scanner only, we leveraged data from $n=175$ individuals who completed both the MRI and behavioral versions of the task ($M=1.75$ years between assessments). First, we fit a linear model predicting behavioral task performance from MRI task performance. The regression weights from this model were then applied to the remaining 485 subjects who only completed the MRI version of the task (e.g., predicted keep track accuracy was $36.0 + 45.7 * \text{MRI accuracy}$). Laboratory and MRI versions of the tasks were highly correlated with one another ($r_s=.52$ to $.88$; see supplement Table S1).

For genetic analyses, we also created residualized versions of all EF and intelligence measures by regressing out age, sex, race (White vs. all other groups), and ethnicity (Hispanic vs. non-Hispanic). Residualization procedures were conducted separately in each subsample (LTS and CAP), then the unstandardized residuals were combined for analysis.

Sample Size and Power Considerations

No *a priori* power calculations were conducted because we planned to include all individuals from the EF-MRI and CATSLife studies with available data at the time of analysis.

Our sample size within LTS had already exceeded that from a previous study when the twins were in adolescence, which demonstrated Common EF and Updating-Specific were significantly genetically correlated with intelligence (Friedman et al., 2008). The addition of 538 individuals from CAP only increased power, especially the adopted sibling pairs (Plomin & DeFries, 1985). This study was not preregistered.

Sample Overlap with Prior Work

Some similar research questions have been examined using earlier data from these subjects. Most notably, Friedman et al. (2008) examined genetic associations between EFs and intelligence using data from only the LTS subjects (i.e., MZ and DZ twins) when they

completed the EF tasks for the first time at mean age 17 years, and Friedman et al. (2006) had reported phenotypic associations with intelligence in an earlier subset of this same dataset. LTS subjects returned at mean age 23 years to complete the EF tasks again. At this wave, we reported on the heritability of EF factors and their associations with personality constructs such as procrastination, impulsivity, and goal management ability (Friedman et al., 2020; Friedman et al., 2016; Gustavson et al., 2015), but not on associations with intelligence. The data examined here represent the complete third and most recent wave of EF data collection for LTS subjects (mean age 29 years). Although the complete age 29 EF results have not been reported in any prior work, data available partway through data collection were included in prior studies. Specifically, phenotypic relationships between EF factor scores and neuroimaging measures were examined in the first 251 individuals tested (Reineberg et al., 2018; Smolker, Friedman, Hewitt, & Banich, 2018), and longitudinal phenotypic and genetic relationships between EFs and stressful life events were examined in the first 587 individuals tested (Morrison et al., 2021).

CAP subjects (i.e., biological and adopted siblings) completed the EF tasks for the first time during CATSLife, so this also represents the first study to examine EFs or their overlap with intelligence. Finally, some publications have reported on intelligence using earlier waves of LTS and/or CAP (Reynolds et al., 2019), including only one preliminary investigation that incorporated CATSLife data collected prior to 2018 (Reynolds et al., 2017). However, none of this work examined associations between EFs and intelligence (aside from the Friedman et al., 2006 and 2008 papers noted above).

Data Analysis

Data screening and trimming followed the same procedures as prior studies with this sample (e.g., Friedman et al., 2008; 2016). RT data excluded errors. Trials after errors were also excluded for the two shifting tasks, as switch vs. repeat trial types are ambiguous if they follow errors. For all RT-based tasks, a within-subject trimming procedure was applied within conditions prior to averaging RTs to obtain the best measures of central tendency (Wilcox & Keselman, 2003).² To reduce the influence of extreme between-subject scores, we replaced task scores farther than three standard deviations from the group mean with values three standard deviations from the mean (<2.5% of observations for any one task). This trimming procedure was conducted separately within LTS-MRI and CATSLife sub-groups (LTS subjects who were not included in the MRI study were included with CAP subjects in the CATSLife sub-group). Scores for the RT measures (i.e., Stroop, both switch costs) were multiplied by -1 in all analyses so higher numbers reflected better performance. Tasks were excluded for participants whose accuracy on that task was not significantly better than chance criteria (<2.5% for any given EF task). When possible, this criterion for each task was determined as the score at which the binomial $p < .01$ for achieving that score by chance.

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. Control conditions are not used for accuracy tasks for which the control conditions (e.g., prosaccade) would be at ceiling.

Phenotypic analyses were conducted using Mplus version 8.3 (Muthén & Muthén, 1998–2017) and genetic analyses were conducted using the OpenMx package in R version 4.1.1 (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 nonsignificant χ^2 or $\chi^2 < \text{two times the degrees of freedom}$, RMSEA $< .06$, and CFI $> .95$ (Hu & Bentler, 1998). 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 et al., 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 the “type=complex” command in Mplus, 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.

Genetic models were based on standard assumptions in twin and adoptive study designs (Neale & Cardon, 1992; Plomin & DeFries, 1983). Variance of a phenotype is separated into proportions attributable to additive genetic influences (A), dominance (non-additive) genetic influences (D), common or shared environmental influences (C), and nonshared environmental influences (E). Additive genetic influences correlate 1.0 in MZ twin pairs, 0.5 in DZ twin pairs and biological siblings, and 0.0 in adoptive siblings because MZ twins share 100% of their alleles identical-by-descent, DZ twins and siblings share (on average) 50% of their alleles identical-by-descent, and adoptive siblings are genetically unrelated. In contrast, dominance genetic influences correlate 1.0 in MZ twin pairs, 0.25 in DZ twin pairs and biological siblings, and 0.0 in adoptive siblings. Common/shared environmental influences (C) correlate 1.0 for all sets of sibling pairs (MZ twins, DZ twins, biological siblings, and adoptive siblings) because they are environmental factors that make siblings more similar to one another. Nonshared environmental influences (E) do not correlate because they are defined as environmental factors that make siblings uncorrelated.

We also assume equal means and variances within pairs and across zygosity. Indeed, in the basic genetic model used to estimate latent variable cross-twin cross-trait correlations (supplement Table S2), we confirmed that we could constrain means and variances for all cognitive tasks across sibling number and across zygosity (MZ/DZ within LTS, biological/adopted pairs within CAP) without a decrement in fit, $\chi^2(84)=85.82$, $p=.424$. The standard assumptions for univariate twin analyses described here also extend to multivariate analyses, including situations where covariances are decomposed into their additive genetic (r_a), dominance (nonadditive) genetic (r_d), shared environmental (r_c), and nonshared environmental components (r_e).

In addition to calculating heritabilities and genetic correlations for additive (a^2 and r_a) and dominance (d^2 and r_d) genetic influences, we also computed broad-sense heritabilities (h^2) and broad-sense genetic correlations (r_g) by summing the contributions of additive and dominance genetic influences within the model (and generating respective 95% CIs).

Specifically, broad-sense heritabilities are computed with: $h^2 = a^2 + d^2$. Broad-sense correlations between trait 1 (T1) and trait 2 (T2) are computed with: $(\text{apath}_{T1} * r_a * \text{apath}_{T2} + \text{dpath}_{T1} * r_d * \text{dpath}_{T2}) / \sqrt{(a^2_{T1} + d^2_{T1}) * (a^2_{T2} + d^2_{T2})}$ where the *apath* and *dpath* are the paths from the additive and dominance latent variables (respectively) to the EF latent variables or full-scale intelligence measure (in Figures 1 and 2). These estimates are interpreted as the total variance (or covariance) explained by all genetic influences and are estimated with much greater precision than the individual additive and dominance estimates.

The latent variable models and univariate heritabilities for individual EF measures were validated in earlier work using the LTS (Friedman et al., 2008). When estimating genetic and environmental correlations between EF factors and intelligence, we did not estimate correlations for variance components estimated at zero, as it does not make sense to calculate a covariance for a variable with no variance (e.g., c^2 has typically been estimated at zero for Updating-specific and Shifting-specific abilities and e^2 has been estimated at zero for Updating-specific ability in previous work using the LTS; Friedman et al., 2016)³. We also removed residual (task-specific) ACE influences that were estimated at zero to aid model convergence for large models.

Finally, because residual variances could not be equated across twin and biological/adoptive sibling subsamples, we did not estimate dominance genetic influences in residual variances for EF tasks. Additionally, we fit two models: (i) a model where all residual variances were freed across subsample (ACE), and (ii) a model where only nonshared environmental residual variances were freed across subsample (E only). For the models focusing on EFs (i.e., not including intelligence), freeing only the E residual variances across subsamples did not fit significantly worse than freeing the A, C, and E residual variances, $\chi^2(12)=10.85$, $p=.540$, indicating that nonshared environmental residual variances were sufficient to account for sample differences at the task level. In the final model with intelligence, we also verified that freeing only the E residual variances did not fit worse than freeing both A and E residual variances, $\chi^2(6)=9.49$, $p=.148$ (all C residual variances had already been removed because they were estimated at zero in the EFs only model).

Data Availability

All stimulus materials and data are available upon request. Additionally, input and output scripts for all phenotypic invariance analyses (in Mplus) and genetic analyses (in R Markdown) are available at the following webpage <https://osf.io/rk6yv/> (Gustavson, Reynolds, & Friedman, 2021).

³In the case of Updating-specific, where the additive genetic variance was also estimated to be $a^2=.00$, we retained both additive and dominance genetic correlations in the subsequent model with intelligence because a dominance-only genetic model is biologically implausible. Moreover, the additive genetic correlation between IQ and Updating-specific ability could not be removed from the model without significant decrease in model fit, $\chi^2(1)=8.03$, $p=.005$.

Results

Descriptive Statistics

Descriptive statistics for all EF tasks and intelligence in both subsamples are displayed in Table 1. Phenotypic correlations among all measures are displayed in Table 2.

Additionally, the supplement displays cross-twin cross-trait correlations between individual study measures (Table S3 and S4) and between EF latent factors and intelligence (Table S2).

Phenotypic Results

EF task intercorrelations were similar in the LTS ($r_s = .12$ to $.53$) and CAP subsamples ($r_s = .11$ to $.52$), as were the correlations between intelligence and individual EF tasks (see Table 2). To confirm whether data from both samples could be collapsed in a single analysis we conducted invariance analyses, displayed in Table 3a (for EF tasks only) and Table 3b (for EF and intelligence measures together). Multi-group models with factor loadings and means constrained across groups (Model 4) had acceptable fit, but residual variances could not be constrained across groups (Model 5). Therefore, we evaluated subsequent genetic models with residual variances freed across subsample. We also estimated a single group model (i.e., with LTS vs. CAP as a covariate; Model 1) which had comparable fit to the best-fitting multi-group models (Model 4).

Results of the best-fitting multi-group model of EFs and intelligence (Model 4 in Table 3b) indicated that intelligence was moderately correlated with the Common EF ($r = .48$, 95% CI [.38, .59], $\chi^2(1) = 83.78$, $p < .001$ for LTS; $r = .50$, 95% CI [.40, .60], $\chi^2(1) = 271.84$, $p < .001$ for CAP) and Updating-specific latent factors ($r = .60$, 95% CI [.48, .72], $\chi^2(1) = 59.28$, $p < .001$ for LTS; $r = .59$, 95% CI [.47, .71], $\chi^2(1) = 122.57$, $p < .001$ for CAP). Intelligence was not significantly correlated with the Shifting-specific latent factor ($r = -.05$, 95% CI [-.15, .05], $\chi^2(1) = 1.02$, $p = .311$ for LTS; $r = -.03$, 95% CI [-.14, .08], $\chi^2(1) = 0.33$, $p = .567$ for CAP).

Most covariates were not systematically associated with study measures in either sample (see Table 2). However, because some associations were observed, all subsequent genetic analyses were based on unstandardized residuals for all EF measures and intelligence after removing variance due to sex, age, race, and ethnicity.

Unity and Diversity Model of Executive Function Measures

The genetic and environmental model of EFs is displayed in Figure 1a, which decomposes total variance in Common EF, Updating-specific, and Shifting-specific factors into additive genetic (A), dominance genetic (D), shared environmental (C), and nonshared environmental (E) variances. Figure 1b displays the same model, except the A and D estimates have been collapsed to show the estimate of the broad-sense heritabilities for each factor (h^2 ; the sum of the additive and dominance genetic influences).

Genetic influences, considering additive and dominance together, explained the majority of the variance in the Common EF (broad-sense heritability; $h^2 = .71$, 95% CI [.52, .94], $\chi^2(2) = 47.50$, $p < .001$), Updating-specific ($h^2 = 1.0$, 95% CI [.65, 1.0], $\chi^2(2) = 23.54$, $p < .001$), and Shifting-specific factors ($h^2 = .61$, 95% CI [.20, .87], $\chi^2(2) = 8.86$, $p = .012$). All EF

factors were explained primarily by dominance genetic influences rather than additive genetic influences, but these estimates were nonsignificant due to low power to differentiate dominance from additive genetic influences (i.e., 95% CIs of broad-sense heritabilities are narrow because they combine both a^2 and d^2 , but individual 95% CIs for a^2 and d^2 are wide for all EF latent factors). Shared environmental influences were nonsignificant for all EF latent factors, but were estimated at nonzero for Common EF ($c^2=.19$, 95% CI [.00, .36]) and Shifting-specific factors ($c^2=.08$, 95% CI [.00, .40]). Indeed, dominance genetic influences, $\chi^2(3)=6.21$, $p=.102$, or shared environmental influences, $\chi^2(3)=2.24$, $p=.524$, on all latent factors could be removed without impacting model fit. Nonshared environmental influences explained 10% of the variance in the Common EF factor ($e^2=.10$, 95% CI [.04, .19], $\chi^2(1)=10.56$, $p=.001$) and 31% of the variance in the Shifting-Specific factor ($e^2=.31$, 95% CI [.13, .52], $\chi^2(1)=12.21$, $p<.001$).

Factor loadings on the three latent EF factors were moderate, and the remaining variance in each EF task was primarily explained by nonshared environmental influences (which includes measurement error at the task level). Some EF tasks had significant residual genetic influences (number–letter and letter memory). Shared environmental influences were estimated at 0 for all EF tasks, so these estimates were fixed to zero in subsequent analyses.

Genetic and Environmental Associations Between Executive Function and Intelligence

The ADCE model of EFs and intelligence is displayed in Figure 2a (see supplement Table S5 for 95% CIs of all estimates). Figure 2b visually displays the same model, but with the additive (A) and dominance (D) genetic influences collapsed into broad-sense heritability ($H = A+D$) and broad-sense genetic correlations (r_g).

There were moderate to strong genetic correlations between intelligence and the Common EF factor (broad-sense genetic correlation; $r_g=.44$, 95% CI [.28, .59], $\chi^2(2)=27.81$, $p<.001$) and the Updating-specific factor ($r_g=.69$, 95% CI [.59, .80], $\chi^2(2)=95.65$, $p<.001$). For Common EF, neither the additive nor dominance genetic correlations were individually significant ($r_a=.29$, 95% CI [−.75, .81], $\chi^2(1)=0.05$, $p=.823$; $r_d=.98$, 95% CI [−1.0, 1.0], $\chi^2(1)=1.79$, $p=.181$). For Updating-specific ability, the genetic correlation with intelligence was driven by additive genetic influences ($r_a=.95$, 95% CI [.56, 1.0], $\chi^2(1)=8.03$, $p=.005$). Consistent with the lack of phenotypic associations, the Shifting-specific factor was not genetically correlated with intelligence (additive, dominance, or broad-sense). In total, genetic influences accounted for 65% of the phenotypic correlation between Common EF and intelligence and 100% of the phenotypic correlation between Updating-specific and intelligence.

Shared environmental correlations were nonsignificant between EF factors and intelligence, $\chi^2(1)<1.41$, $p>.235$. The nonshared environmental influences on intelligence were correlated with those for Common EF ($r_e=.60$, 95% CI [.33, .94], $\chi^2(1)=19.26$, $p<.001$), explaining 16% of their phenotypic correlation, but were not correlated with the other EF factors.

Finally, after accounting for genetic associations with the EF factors, the remaining variance in intelligence was estimated to be zero for additive genetic influences ($a^2 = .00$, 95%

CI [.00, .30]), dominance genetic influences ($d^2 = .00$, 95% CI [.00, .16]), and shared environmental influences ($c^2 = .00$, 95% CI [.00, .12]). Nonshared environmental influences unique to intelligence remained statistically significant, ($e^2 = .10$, 95% CI [.01, .15], $\chi^2(1)=4.20$, $p=.040$).

Discussion

This study examined the genetic and environmental overlap among 3 EF latent factors and full-scale intelligence in a combined sample of twins, biological siblings, and unrelated adopted siblings. This was the first genetic study of Common EF of adults in this age range (late 20s to early 40s) and the first genetic study of Common EF to utilize an extended family design. Results confirmed that both Common EF and Updating-specific ability were moderately to strongly correlated with intelligence, whereas Shifting-specific ability was unrelated to intelligence.

The genetic design revealed multiple findings relevant to our understanding of EFs, intelligence, and their overlap. First, associations between intelligence and Common EF were explained by both genetic influences (total broad-sense genetic correlation; $r_g=.44$) and nonshared environmental influences ($r_e=.60$). The broad-sense genetic correlation ($r_g=.44$, 95% CI [.28, .59]) was similar in magnitude to the additive genetic correlations observed in older adolescents (age 17; $r_a=.57$; CIs not reported; Friedman et al., 2008) and middle-aged adults (age 56; $r_a=.59$, 95% CI [.56, 1.00]; Gustavson et al., 2018b), but was smaller than those observed in younger adolescents (age 10: $r_a=.86$; 95% CI [.45, 1.27]; Freis et al, 2021; age 11: $r_a=.92$; 95% CI [.72, 1.11]; Englehardt et al. 2016). Even with the extended family design employed here, there was still low power to detect whether the strong genetic correlation was driven by additive genetic influences and dominance genetic influences. However, consistent with twin correlations from earlier work (e.g., Friedman et al., 2008), and with earlier extended family approaches (Keller et al., 2013), there was little evidence for nonadditive genetic influences on intelligence ($d^2=.11$), whereas the dominance genetic influences estimated for Common EF ($d^2=.65$), updating-specific ($d^2=.45$), and shifting-specific ($d^2=.60$) were nominally higher, suggesting a larger role of nonadditive genetic influences on EFs than intelligence that should be explored using larger datasets.

Second, all factors (except Updating-specific) had some estimated shared environmental influences ($c^2=.10$ to .18). These estimates were not significant in this sample, but they were similar in magnitude to estimates from earlier studies of adults for intelligence (e.g., Keller et al., 2013; Lyons et al., 2017). Additional analyses revealed that removing dominance influences from the model (i.e., ACE model) reduced the shared environmental estimate on Common EF ($c^2=.18$ in Figure 2 vs. $c^2=.03$ in Figure S1). This pattern suggests that dominance genetic influences may mask the contribution of shared environment influences on EFs in the standard twin design, although that dominance variance was not statistically significant.

Third, the phenotypic correlation between Updating-specific ability and intelligence was driven entirely by genetic influences (broad-sense $r_g=.69$). These findings suggest that intelligence is strongly genetically associated with Updating-specific ability in addition to

Common EF ability, consistent with decades-old proposals regarding the strong link between intelligence and working memory processes (Colom et al., 2004; Kyllonen & Christal, 1990). Recall that performance on any given working memory updating task is influenced by both Common EF and Updating-specific processes. These findings demonstrate that working memory tasks are most strongly associated with intelligence because both processes are independently genetically correlated with intelligence (with an environmental correlation through Common EF as well). In fact, in these data, there were no genetic influences remaining for intelligence after accounting for both Common EF and Updating-specific abilities. Therefore, although both EF abilities are genetically distinct from intelligence, when combined they may explain all the genetic variance in intelligence, with only some (nonsignificant) residual shared and nonshared environmental influences on intelligence.

The above conclusions would account for the fact that phenotypic and genetic associations between Common EF and intelligence approach unity in childhood and early adolescence (Brydges et al., 2012; Engelhardt et al., 2016; Freis et al., 2021) but weaken somewhat in later adolescence and adulthood (Friedman et al., 2008; Gustavson et al., 2018b). The factor structure of EFs appears to be unitary in children before about age 10–12 (Brydges et al., 2014; Hartung et al., 2020; Xu et al., 2013). By late adolescence and into adulthood, however, the genetic correlation between Updating-specific abilities and intelligence can be observed in addition to the genetic correlation between common EF and intelligence (Friedman et al., 2008). It will be interesting to see if a similar pattern emerges in old age, as some have argued that EF processes may again converge in old adults, especially cognitively impaired older adults (de Frias, Dixon, & Strauss, 2009). If this is the case, Common EF and intelligence may again be difficult to disentangle (especially genetically) in older and/or impaired samples.

Figure 3 summarizes existing studies that have examined the heritability of Common EF, including our new results. Our findings fit between those from LTS subjects at mean age 23 (Friedman et al., 2016) and those of middle-aged adults in the Vietnam Era Twin Study of Aging at mean age 56 (Gustavson et al., 2018a). Shared environmental influences appear to be re-emerging as individuals approach middle age, and nonshared environmental influences appear to stay relatively stable throughout young adulthood into midlife. However, as noted earlier, an alternative interpretation is that some shared environmental influences exist in adolescence and young adulthood, but are masked by the presence of dominance genetic variance that cannot be simultaneously estimated with shared environmental influences in a standard twin design.

Still other factors besides age may have impacted our estimates of the genetic and environmental influences on EF and their overlap with intelligence compared to earlier work. Results may have been affected by sampling differences (e.g., socioeconomic status), including those differences between CAP and LTS described in the limitations section. They may also be explained by the fact that we administered only two tasks per EF domain, whereas more measures were administered in adolescent samples (Engelhardt et al., 2016) including our earlier investigation of the LTS sample (Friedman et al., 2008). Results may also be impacted by modeling differences, though there was little evidence that our choice

of using full-scale IQ instead of a latent variable for general cognitive ability impacted the results (see supplement Figure S2).

Shifting-specific ability was not correlated with intelligence phenotypically or at the genetic/environmental levels. Previously, Friedman et al., (2008) found a small negative correlation between intelligence and Shifting-specific ability in the LTS sample assessed in late adolescence. This negative association was interpreted as reflecting a stability-flexibility tradeoff (Goschke, 2000; Miyake & Friedman, 2012), wherein individuals that are able to strongly engage a task-set representation may have difficulty flexibly switching away to new task-sets. Our findings do not rule out a weak negative correlation between Shifting-specific ability and intelligence but are more consistent with intelligence being simply uncorrelated with Shifting-specific ability in adulthood. Nevertheless, because intelligence measures are typically positively correlated with all other cognitive abilities, it is interesting that it may be entirely unrelated to Shifting-specific ability, potentially reflecting entirely separable brain circuitry.

Strengths and Limitations

This study examined EFs and intelligence data from combined twin and adoptive/biological sibling datasets, greatly increasing power to detect shared environmental influences compared to traditional twin methods (Plomin & DeFries, 1985). However, even with this large sample we did not have sufficient power to distinguish between additive and dominance (nonadditive) genetic influences in many analyses (Martin et al., 1978). Twin correlations for EF tasks often suggest dominance (within-pair $r_{MZ} > 2 * r_{DZ}$), whereas twin correlations for intelligence typically suggest (weak) shared environmental influences (within-pair $r_{MZ} < 2 * r_{DZ}$). Modeling additive and dominance genetic influences alongside shared and nonshared environmental influences using the extended family design is therefore useful in obtaining unbiased estimates (e.g., for investigation in larger samples).

The CAP sample was also slightly older, which may have contributed to the fact that we could not collapse residual variances across subsamples. Moreover, both samples were predominantly White, which impacts the generalizability of findings. However, both samples were representative of the Colorado population at the time of recruitment. Additionally, although some assumptions of the twin/family design were supported (e.g., equal means and variance across sibling number and zygosity within each sample), violation of other assumptions could influence our results. Specifically, the same-sex, same-aged twins from LTS may share larger proportions of their shared environment than biological or adopted siblings from CAP.

CAP subjects also completed all EF tests in the laboratory, whereas most LTS subjects completed 3 of the 6 EF tasks in the scanner. However, the remaining 3 tasks were completed in the same laboratory setting as the remaining subjects, and we were able to harmonize the 3 scanner tasks using data from the $n=175$ subjects who completed both sets of tasks. Indeed, reanalysis excluding data from the 3 MRI tasks revealed similar patterns of results to those displayed in Figure 2 (see supplemental Figure S3).

Genetic and environmental variances were obtained by estimating paths on latent ADCE factors (i.e., estimating the square roots of the variances). This model specification means that variance components cannot be negative. Although a negative variance component is not interpretable, it is possible for a parameter whose true value is zero to be negative due to sampling error, and imposing a boundary on this parameter could lead to biased estimates for remaining parameters (Verhulst et al., 2019). However, the likelihood-based confidence intervals in OpenMx include an adjustment for parameters that are estimated close to a boundary.

Finally, our analyses of the overlap between EF and intelligence was based on the WAIS-III full-scale IQ score, which combines across measures of verbal and performance intelligence. Full scale IQ (or equivalent combined scores) are the most widely used in this research area (Engelhardt et al., 2016; Friedman et al., 2008; Gustavson et al., 2018b), and prior work with the LTS sample suggests that the associations of EFs with full scale IQ are highly similar to those with a general factor extracted from the WAIS subtests (Friedman et al., 2008). Indeed, in the current sample, we found similar results when modeling intelligence as a latent factor with the verbal and performance intelligence scores as indicators (supplement Figure S2). Prior work with the LTS sample also suggests that Inhibiting, Updating, and Shifting latent variables show similar relationships to latent factors for fluid and crystallized intelligence in this neuropsychologically intact sample (Friedman et al., 2006; Salthouse et al., 2003; Unsworth et al., 2009), although associations with EFs may differ across these different aspects of intelligence in samples with brain lesions (Duncan et al., 1995; Duncan et al., 1996).

Summary and Conclusions

For decades, researchers have debated whether EFs and intelligence truly represent distinct constructs. Using a combined twin and extended family design of adults, our data suggest that intelligence is distinct from both Common EF and Updating-specific abilities, but that together these two components of EF may account for all of the genetic influences on intelligence. Intelligence and EFs may be further distinguished by different patterns of genetic and environmental influences: Intelligence has stronger evidence for shared environmental variance whereas EFs have stronger evidence for dominance genetic variance, though a larger extended twin and family design would be needed to estimate these variance components with precision. These findings highlight the utility of genetic designs in better understanding the structure and separability of cognitive abilities, as the association between EFs and intelligence at the genetic level is much stronger than that at the environmental level. They also highlight the utility of the unity and diversity model of EF as a way of understanding why working memory processes have historically been closely related to (and difficult to distinguish from) intelligence. Such working memory processes reflect a combination of Common EF and Updating-specific processes, which are independently moderately related to intelligence.

Context

The unity and diversity model of executive functions (EFs), which posits that EFs can be delineated into general (Common EF) and specific (Updating-specific, Shifting-specific)

abilities, was developed in adolescents and validated in studies of children, young adults, and middle aged adults. The separability of EFs (particularly Common EF) from intelligence continues to be debated in the field. The authors saw an opportunity to examine these associations in an adult sample (late 20s to early 40s), which both (i) fills a gap in the age demographic of existing studies of EFs, and (ii) allows for evaluation of the hypothesis that intelligence is separately genetically correlated with Common EF and Updating-specific abilities, which may only be sufficiently separable constructs starting in adolescence. Furthermore, the availability of biological and adopted siblings allowed for a genetic investigation of EFs and intelligence in a sample that did not rely solely on twins. Results are consistent with the hypothesis that EFs may also differ from intelligence by being more strongly influenced by dominance genetic variance, and that intelligence is related to both Common EF and Updating-Specific abilities. Together, these results suggest that although they share genetic and environmental influences, EFs and intelligence have different genetic/environmental structures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This work was supported by supported by NIH grants R03AG065643, R01AG046938, R01DC016977, R01MH063207, R01DA046064, and U01DA046413. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Brydges CR, Fox AM, Reid CL, & Anderson M (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence*, 47, 34–43. doi:10.1016/j.intell.2014.08.010
- Brydges CR, Reid CL, Fox AM, & Anderson M (2012). A unitary executive function predicts intelligence in children. *Intelligence*, 40, 458–469. doi:10.1016/j.intell.2012.05.006
- Carpenter PA, Just MA, & Shell P (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404–431. doi:10.1037/0033-295x.97.3.404 [PubMed: 2381998]
- Colom R, Rebollo I, Palacios A, & Juan-Espinosa MK, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, 32, 277–296. doi:10.1016/j.intell.2003.12.002
- Corley RP, Reynolds CA, Wadsworth SJ, Rhea SA, & Hewitt JK (2019). The Colorado Twin Registry: 2019 update. *Twin Research and Human Genetics*, 22, 707–715. doi:10.1017/thg.2019.50 [PubMed: 31708007]
- de Frias CM, Dixon RA, & Strauss E (2009). Characterizing executive functioning in older special populations: From cognitively elite to cognitively impaired. *Neuropsychology*, 23, 778–791. doi:10.1037/a0016743 [PubMed: 19899836]
- Dempster FN (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence*, 15, 157–173. doi:10.1016/0160-2896(91)90028-C
- Duncan J, Burgess P, & Emslie H (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, 33, 261–268. doi:10.1016/0028-3932(94)00124-8 [PubMed: 7791994]
- Duncan J, Emslie H, Williams P, Johnson R, & Freer C (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognitive Psychology*, 30, 257–303. doi:10.1006/cogp.1996.0008 [PubMed: 8660786]

- 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]
- Engle RW, Tuholski SW, Laughlin JE, & Conway ARA (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128, 309–331. doi:10.1037/0096-3445.128.3.309 [PubMed: 10513398]
- Freis SM, Morrison CL, Lessem JM, Hewitt JK, & Friedman NP (2021). Genetic and environmental influences on executive functions and intelligence in middle childhood. *Developmental Science*, e13150. doi:10.1111/desc.13150 [PubMed: 34288270]
- Friedman NP, du Pont A, Corley RP, & Hewitt JK (2018). Longitudinal relations between depressive symptoms and executive functions from adolescence to early adulthood: A twin study. *Clinical Psychological Science*, 6, 543–560. doi:10.1177/2167702618766360 [PubMed: 30250762]
- Friedman NP, Hatoum AS, Gustavson DE, Corley RP, Hewitt JK, & Young SE (2020). Executive functions and impulsivity are genetically distinct and independently predict psychopathology: Results from two adult twin studies. *Clinical Psychological Science*, 8, 519–538. doi:10.1177/2167702619898814 [PubMed: 33758683]
- 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, Corley RP, Young SE, Defries JC, & Hewitt JK (2006). Not all executive functions are related to intelligence. *Psychological Science*, 17, 172–179. doi:10.1111/j.1467-9280.2006.01681.x [PubMed: 16466426]
- 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]
- Fujisawa KK, Todo N, & Ando J (2017). Genetic and environmental influences on the development and stability of executive functions in children of preschool age: A longitudinal study of Japanese twins. *Infant and Child Development*, 26, e1994. doi:10.1002/icd.1994
- Fujisawa KK, Todo N, & Ando J (2019). Changes in genetic and environmental influences on cognitive ability, executive function, and preacademic skills in Japanese preschool age twins. *Developmental Psychology*, 55, 38–52. doi:10.1037/dev0000627 [PubMed: 30359061]
- 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, Franz CE, Panizzon MS, Reynolds CA, Xian H, Jacobson KC, ... Kremen WS (2019). Genetic and environmental associations among executive functions, trait anxiety, and depression symptoms in middle age. *Clinical Psychological Science*, 7, 127–142. doi:10.1177/2167702618805075 [PubMed: 30923645]
- Gustavson DE, Miyake A, Hewitt JK, & Friedman NP (2015). Understanding the cognitive and genetic underpinnings of procrastination: Evidence for shared genetic influences with goal management and executive function abilities. *Journal of Experimental Psychology: General*, 144, 1063–1079. doi:10.1037/xge0000110 [PubMed: 26389573]
- Gustavson DE, Panizzon MS, Elman JA, Franz CE, Reynolds CA, Jacobson KC, ... Kremen WS (2018a). 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 (2018b). Genetic and environmental architecture of executive functions in midlife. *Neuropsychology*, 32, 18–30. doi:10.1037/neu0000389 [PubMed: 28857599]
- Gustavson DE, Reynolds CA, & Friedman NP (2021). Genetic associations between executive functions and intelligence: A combined twin and adoption study. Retrieved from osf.io/rk6yv
- Hartung J, Engelhardt LE, Thibodeaux ML, Harden KP, & Tucker-Drob EM (2020). Developmental transformations in the structure of executive functions. *Journal of Experimental Child Psychology*, 189, 104681. doi:10.1016/j.jecp.2019.104681 [PubMed: 31648081]
- 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]
- Keller MC, Garver-Apgar CE, Wright MJ, Martin NG, Corley RP, Stallings MC, ... Zietsch BP (2013). The genetic correlation between height and IQ: Shared genes or assortative mating? *PLoS Genet*, 9, e1003451. doi:10.1371/journal.pgen.1003451 [PubMed: 23593038]
- Kyllonen PC, & Christal RE (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14, 389–433. doi:10.1016/s0160-2896(05)80012-1
- 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*, 40, 97–116. doi:10.1038/hdy.1978.10 [PubMed: 272366]
- Matteson LK, McGue M, & Iacono WG (2013). Shared environmental influences on personality: A combined twin and adoption approach. *Behavior Genetics*, 43, 491–504. doi:10.1007/s10519-013-9616-8 [PubMed: 24065564]
- Mayr U, & Kliegl R (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1124–1140. doi:10.1037//0278-7393.26.5.1124 [PubMed: 11009248]
- 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/cogn.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
- Morrison CL, Rhee SH, Smolker HR, Corley RP, Hewitt JK, & Friedman NP (2021). Genetic and environmental influences on stressful life events and their associations with executive functions in young adulthood: A longitudinal twin analysis. *Behavior Genetics*, 51, 30–44. doi:10.1007/s10519-020-10017-9 [PubMed: 32959091]
- Muthén LK, & Muthén BO (1998–2017). *Mplus User’s Guide: Eighth Edition*. 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]
- Neisser U, Boodoo G, Bouchard TJ Jr, Boykin AW, Brody N, Ceci SJ, ... Sternberg RJ (1996). Intelligence: Knowns and unknowns. *American psychologist*, 51, 77.
- Plomin R, & DeFries JC (1983). The Colorado Adoption Project. *Child Development*, 54, 276–289. doi:10.2307/1129691 [PubMed: 6872628]
- Plomin R, & DeFries JC (1985). The origins of individual differences in infancy; the Colorado adoption project. *Science*, 230, 1369–1371. [PubMed: 17749676]
- 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:10.1375/twin.9.3.367 [PubMed: 16790147]
- Reineberg AE, Gustavson DE, Benca CE, Banich MT, & Friedman NP (2018). The relationship between resting state network connectivity and individual differences in executive function. *Frontiers in Psychology*, 9. doi:10.3389/fpsyg.2018.01600
- Reynolds CA, Smolen A, Corley RP, Munoz E, Friedman NP, Rhee SH, ... Wadsworth SJ (2019). APOE effects on cognition from childhood to adolescence. *Neurobiology of Aging*, 84, 239 e231–239 e238. doi:10.1016/j.neurobiolaging.2019.04.011
- Reynolds CA, Smolen A, Haberstick BC, DeFries JC, & Wadsworth SJ (2017). APOE effects on cognition from middle childhood to the cusp of middle adulthood. *Innovation in Aging*, 1, 475. doi:10.1093/geroni/igx004.1690
- 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
- Rowe DC, Vesterdal WJ, & Rodgers JL (1998). Herrnstein's syllogism: Genetic and shared environmental influences on IQ, education, and income. *Intelligence*, 26, 405–423. doi:10.1016/s0160-2896(99)00008-2
- Salthouse TA (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology*, 19, 532–545. doi:10.1037/0894-4105.19.4.532 [PubMed: 16060828]
- Salthouse TA, Atkinson TM, & Berish DE (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, 132, 566–594. doi:10.1037/0096-3445.132.4.566 [PubMed: 14640849]
- Satorra A, & Bentler PM (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507–514. doi:10.1007/Bf02296192
- Smolker HR, Friedman NP, Hewitt JK, & Banich MT (2018). Neuroanatomical correlates of the unity and diversity model of executive function in young adults. *Frontiers in Human Neuroscience*, 12, 283. doi:10.3389/fnhum.2018.00283
- Spearman C (1904). "General Intelligence" Objectively Determined and Measured. *American Journal of Psychology*, 15, 201–293. doi:10.1037/11491-006
- 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
- Unsworth N, Miller JD, Lakey CE, Young DL, Meeks JT, Campbell WK, & Goodie AS (2009). Exploring the relations among executive functions, fluid intelligence, and personality. *Journal of Individual Differences*, 30, 194–200. doi:10.1027/1614-0001.30.4.194
- Verhulst B, Prom-Wormley E, Keller M, Medland S, & Neale MC (2019). Type I error rates and parameter bias in multivariate behavioral genetic models. *Behavior Genetics*, 49, 99–111. doi:10.1007/s10519-018-9942-y [PubMed: 30569348]

- Wadsworth SJ, Corley RP, Munoz E, Trubenstein BP, Knaap E, DeFries JC, ... Team CA (2019). CATSLife: A study of lifespan behavioral development and cognitive functioning. *Twin Research and Human Genetics*, 1–12. doi:10.1017/thg.2019.49 [PubMed: 30661510]
- Wechsler D (1997). *Wechsler adult intelligence scale* (3rd ed.). San Antonio, TX: Psychological Corporation.
- 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]
- Xu F, Han Y, Sabbagh MA, Wang T, Ren X, & Li C (2013). Developmental differences in the structure of executive function in middle childhood and adolescence. *PLoS One*, 8, e77770. doi:10.1371/journal.pone.0077770 [PubMed: 24204957]
- Yntema DB (1963). Keeping track of several things at once. *Human Factors*, 5, 7–17. doi:10.1177/001872086300500102 [PubMed: 14002586]

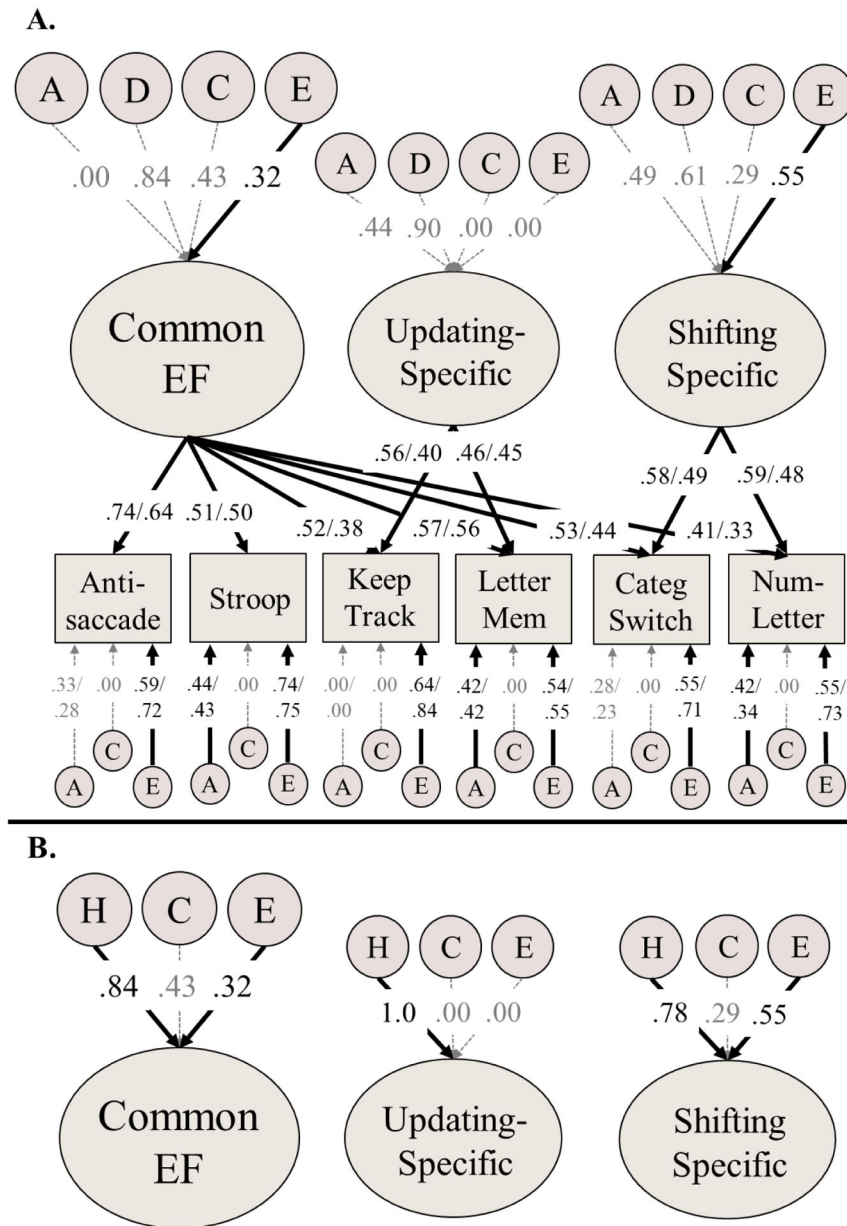


Figure 1: Genetic model of executive functions (EFs). Panel B displays the same model, but results are plotted for broad-sense heritabilities (H = additive + dominance genetic variance). Variance explained by latent variables (ovals) can be computed by squaring the factor loadings on measured variables (rectangles) or EF subfactors. Most paths were equated across subsample, but because residual nonshared environmental influences were allowed to differ, standardized estimates of factor loadings and residual A influences also differ across Longitudinal Twin Study subsample (LTS; left) and the Colorado Adoption Project subsample (CAP; right). Significant paths are displayed in black text, with solid black arrows ($p < .05$). Model fit: $\chi^2(307)=464.09$, $p < .001$, RMSEA=.026, CFI=.912. A =

Additive genetic influences; C = Shared/common environmental influences; E = Nonshared environmental influences.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

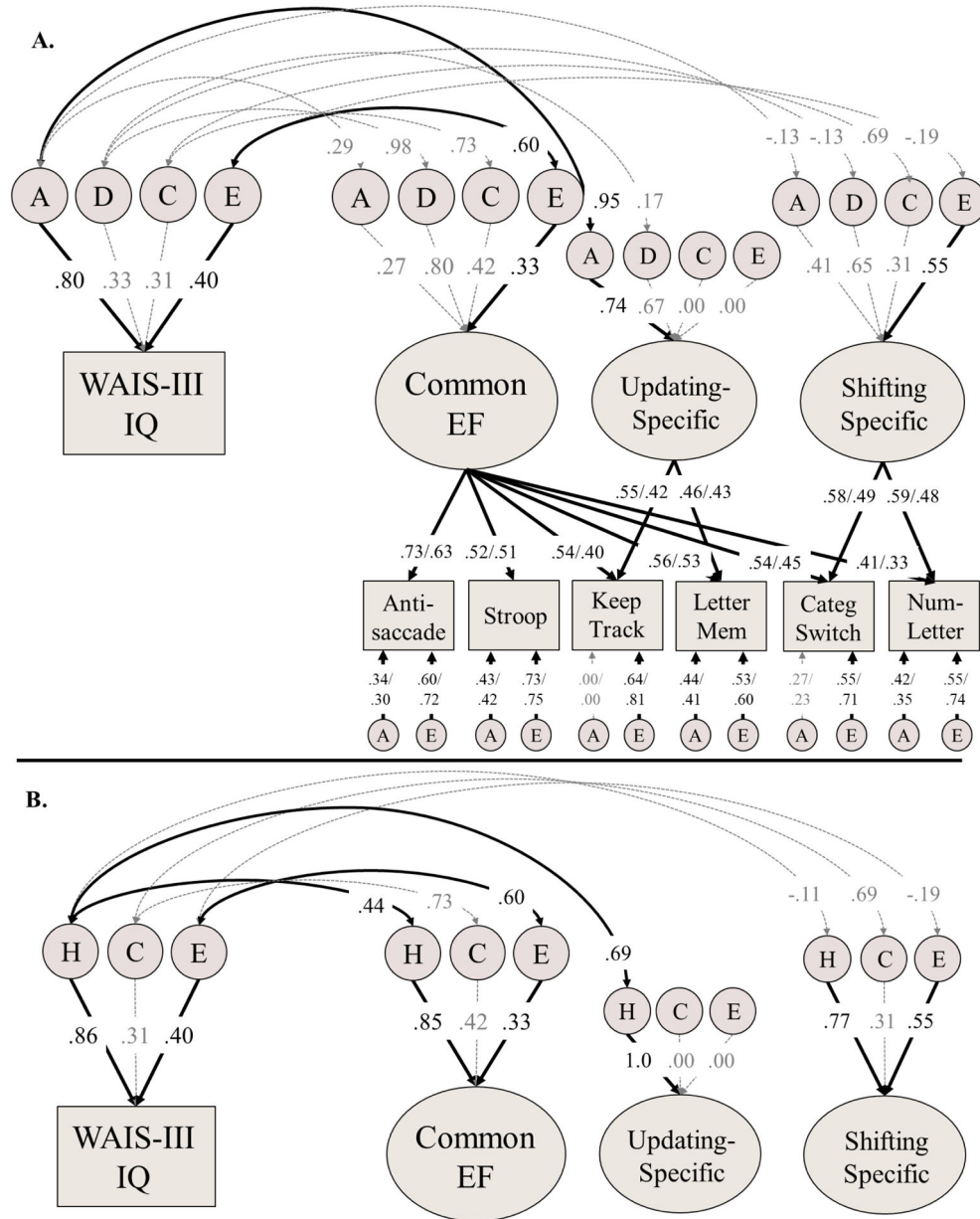
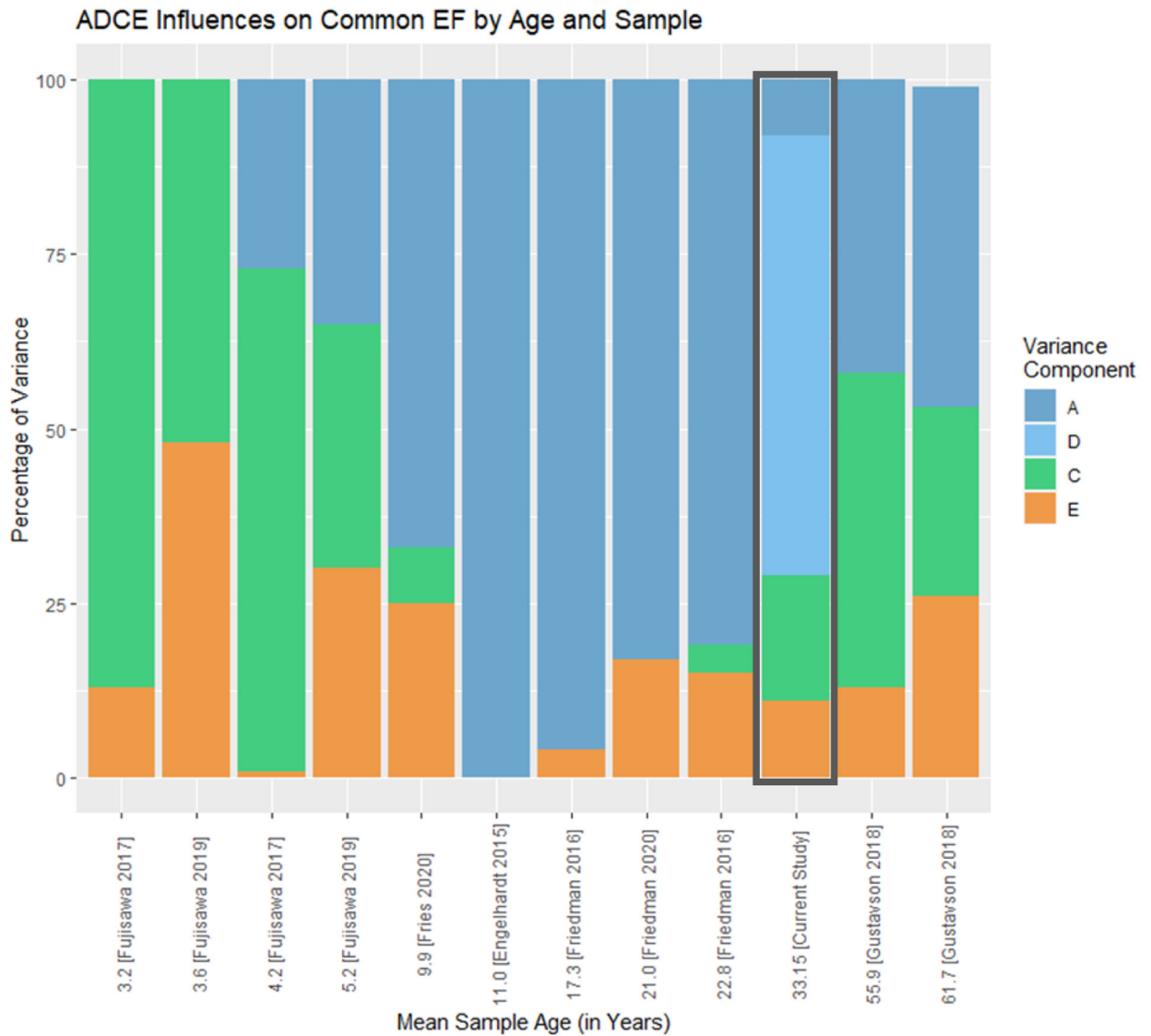


Figure 2: Genetic model of executive functions (EFs) and intelligence. Panel B displays the same model, but results are plotted for broad-sense heritabilities (H = additive + dominance genetic variance) and broad-sense genetic correlations (r_g). Variance explained by latent variables (ovals) can be computed by squaring the factor loadings on measured variables (rectangles) or EF latent factors. Genetic and environmental correlations with IQ were computed from the statistically-equivalent Cholesky decomposition. Most paths were equated across subsample, but because residual nonshared environmental influences were allowed to differ, standardized estimates of factor loadings and residual A influences also differ across Longitudinal Twin Study subsample (LTS; left) and the Colorado Adoption Project subsample (CAP; right). Significant paths are displayed in black text, with solid

black arrows ($p < .05$). See supplement Table S5 for 95% confidence intervals for this model. Model fit: $\chi^2(413)=659.02$, $p < .001$, RMSEA=.029, CFI=.907. A = Additive genetic influences; D = Dominance (nonadditive) genetic influences; C = Shared/common environmental influences; E = Nonshared environmental influences; WAIS-III IQ = Wechsler Adult Intelligence Scale full-scale intelligence quotient.

**Figure 3:**

Summary of additive genetic (A), dominance genetic (D), shared environmental (C) and nonshared environmental (E) influences on Common EF based on age of assessment across existing studies. All studies assessed Common EF using latent factors (or a composite of 5 to 9 tasks in the case of the Fujisawa et al., 2017 & 2019 studies). The current study (outlined with a grey box) is also the only study to estimate D influences. Therefore, C estimates could be suppressed in other studies. Individual EF tasks differed across samples; thus, it is also possible that modeling and measurement differences contribute to the variability in ADCE estimates. However, there is some fluctuation in ACE across samples that are overlapping in subjects and measures. Subjects from Fujisawa, Todo, and Ando (2017) and Fujisawa, Todo, and Ando (2019) may be overlapping (Tokyo Twin Cohort Project). Subjects are also overlapping within Friedman et al. (2016) (Colorado Longitudinal Twin Study) and within Gustavson et al. (2018a) (Vietnam Era Twin Study of Aging). Data from

the current project is partially overlapping with Friedman et al. (2016) (LTS, but not CAP subjects).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Descriptive Statistics for All Study Measures

Measure	N	M	SD	Range	Skewness	Kurtosis	Reliability
<i>Longitudinal Twin Study</i>							
Antisaccade (% correct) - Laboratory	228	72.68	17.38	16.67, 98.61	-0.62	-0.25	0.91 ^a
Antisaccade (% correct) - MRI*	448	-0.02	0.99	-2.36, 2.33	-0.17	-0.73	0.94 ^a
Stroop (RT Interference)	696	151.35	78.66	-90.29, 426.38	0.82	1.12	0.96 ^a
Number Letter (RT Switch Cost) - Laboratory	231	260.21	193.19	-81.90, 928.69	1.21	1.66	0.91 ^a
Number Letter (RT Switch Cost) - MRI	451	180.00	117.94	-29.58, 565.54	0.86	0.85	0.93 ^a
Category Switch (RT Switch Cost)	692	184.03	151.63	-103.05, 800.11	1.14	1.46	0.91 ^a
Keep Track (% correct) - Laboratory	234	70.19	11.93	26.63, 97.22	-0.71	0.88	0.64 ^b
Keep Track (% correct) - MRI	462	76.50	14.01	34.96, 100.00	-0.68	0.13	0.75 ^b
Letter Memory (% correct)	697	72.77	14.18	25.73, 100.00	-0.09	-0.73	0.93 ^b
WAIS Full Scale IQ	684	108.87	11.31	74, 143	-0.14	-0.01	0.97 ^c
<i>Colorado Adoption Project</i>							
Antisaccade (% correct)	533	66.12	17.43	15.28, 98.61	-0.37	-0.55	0.90 ^a
Stroop (RT Interference)	537	176.20	82.12	1.93, 426.38	0.70	0.61	0.96 ^a
Number Letter (RT Switch Cost)	534	294.83	194.20	-157.13, 928.69	0.98	1.39	0.91 ^a
Category Switch (RT Switch Cost)	537	252.20	188.32	-163.94, 808.37	0.94	0.66	0.93 ^a
Keep Track (% correct)	534	65.75	13.61	26.63, 94.44	-0.48	0.01	0.60 ^b
Letter Memory (% correct)	527	69.88	15.06	25.73, 100.00	-0.14	-0.38	0.93 ^b
WAIS Full Scale IQ	534	112.42	12.17	69, 148	-0.07	0.18	0.97 ^c

Note: Reliability is split-half Spearman-Brown coefficient (^a), Cronbach's alpha (^b), or internal reliability from Wechsler (1997) (^c). Reliability for MRI subjects was based on all subjects who completed that version, including the $n=153$ individuals whose primary data was analyzed based on their CATSLife assessment. MRI scores are shown prior to harmonization. See supplement Table S6 for descriptive statistics of harmonized MRI measures and for statistics collapsing across laboratory and MRI versions.

* MRI Antisaccade scores were z-scored because one of the cue durations was changed from 233ms to 283 ms in the middle of the study (z-scoring was done within each cue duration group, then concatenated). RT = Reaction time.; WAIS = Wechsler Adult Intelligence Scale.

Table 2

Correlations Among All Study Measures

Measure	<i>N</i>	1	2	3	4	5	6	7
<i>Longitudinal Twin Study</i>								
1. Antisaccade (% correct)	676	1						
2. Stroop (RT Interference)	696	0.24	1					
3. Number Letter (RT Switch Cost)	682	0.29	0.26	1				
4. Category Switch (RT Switch Cost)	692	0.38	0.31	0.53	1			
5. Keep Track (% correct)	696	0.34	0.24	0.15	0.23	1		
6. Letter Memory (% correct)	697	0.40	0.32	0.12	0.23	0.51	1	
7. WAIS Full-Scale IQ	684	0.32	0.26	0.07	0.22	0.53	0.55	1
8. Sex (Female / Male)	700	0.18	-0.11	-0.05	0.05	-0.08	-0.03	0.03
9. Age (at EF assessment)	700	0.04	-0.08	-0.05	-0.08	-0.07	-0.11	-0.02
10. Race (Non-White / White)	700	0.00	-0.02	0.04	0.05	-0.10	-0.08	-0.11
11. Ethnicity (Non-Hispanic / Hispanic)	700	0.02	0.03	0.05	0.06	-0.05	-0.06	-0.14
<i>Colorado Adoption Project</i>								
1. Antisaccade (% correct)	533	1						
2. Stroop (RT Interference)	537	0.33	1					
3. Number Letter (RT Switch Cost)	534	0.22	0.25	1				
4. Category Switch (RT Switch Cost)	537	0.32	0.31	0.44	1			
5. Keep Track (% correct)	534	0.31	0.26	0.13	0.18	1		
6. Letter Memory (% correct)	527	0.39	0.34	0.11	0.14	0.52	1	
7. WAIS Full-Scale IQ	534	0.36	0.31	0.06	0.21	0.61	0.49	1
8. Sex (Female / Male)	538	0.26	0.03	-0.11	0.03	-0.05	0.02	0.13
9. Age	538	-0.07	0.04	-0.11	-0.10	-0.04	0.04	0.03
10. Race (Non-White / White)	538	-0.04	-0.01	0.01	0.00	-0.06	-0.01	-0.06
11. Ethnicity (Non-Hispanic / Hispanic)	538	0.06	0.00	0.00	-0.02	0.02	0.09	0.02

Note: Stroop and switch cost scores based on reaction times (RTs) were reverse scored prior to analyses so that higher numbers indicate better performance on all measures.

Table 3

Tests of Invariance Across LTS and CAP Subsamples

Model	χ^2	df	<i>p</i>	RMSEA	CFI
A. EF data only					
<i>Single-Group Model</i>					
1. LTS vs. CAP as a covariate	49.22	7	<.001	0.070	0.971
<i>Multi-Group Models</i>					
2. No constraints	57.25	14	<.001	0.071	0.968
3. Constrain factor loadings across groups	62.11	19	<.001	0.061	0.968
4. Model 3 + constrain intercepts	78.78	22	<.001	0.065	0.960
5. Model 4 + constrain residual variances	236.65	31	<.001	0.104	0.847
6. Model 5 + constrain latent variable variances	299.21	34	<.001	0.112	0.903
B. EF and Intelligence					
<i>Single Group Model</i>					
1. LTS vs. CAP sample as a covariate	67.31	10	<.001	0.068	0.972
<i>Multi-Group Models</i>					
2. No constraints	109.01	20	<.001	0.085	0.954
3. Constrain factor loadings across groups	118.93	25	<.001	0.078	0.952
4. Model 3 + constrain intercepts	137.96	32	<.001	0.073	0.946
5. Model 4 + constrain residual variances	268.44	38	<.001	0.099	0.882
6. Model 5 + constrain latent variable variances	343.71	42	<.001	0.108	0.846

Note: All models included regression terms for age, sex, race (White vs. Non-White), and ethnicity (Hispanic vs. Non-Hispanic) separately for each subsample. The models in bold indicate acceptable fit after constraining factor loadings and intercepts across the Longitudinal Twin Study (LTS) and Colorado Adoption Project (CAP) subsamples, indicating they could be collapsed for genetic analyses. EF = Executive function.