Stability of Genetic and Environmental Influences on Executive Functions in Midlife

Daniel E. Gustavson[1], **Matthew S. Panizzon**[1], **Jeremy A. Elman**[1], **Carol E. Franz**[1], **Chandra A. Reynolds**[2], **Kristen C. Jacobson**[3], **Naomi P. Friedman**[4], **Hong Xian**[5], **Rosemary Toomey**[6], **Michael J. Lyons**[6], and **William S. Kremen**[1]

[1]University of California, San Diego

[2]University of California, Riverside

[3]University of Chicago

[4]University of Colorado Boulder

[5]St. Louis University

[6]Boston University

Abstract

Research on executive functions (EFs) has revealed that individual differences in general EF abilities are highly correlated across the first few decades of life, especially at the level of genetic influences. Our work has also provided evidence for substantial heritability of this Common EF factor in midlife, but it remains unclear whether individual differences in Common EF continue to show strong stability in middle age. We examined data from 1464 middle-aged twins from the Vietnam Era Twin Study of Aging, most of whom completed seven neuropsychological measures of EFs at two points in middle age (mean ages 56 and 62). Confirmatory factor analysis indicated that individual differences in Common EF, a latent factor explaining variation in seven neuropsychological EF tasks, were highly correlated across this six-year period ($r = .97$), and that the same genetic and environmental influences were operating across this interval (genetic and shared environmental correlations = 1.0, nonshared environment correlation = .95). Similar phenotypic and genetic stability was observed for a Working Memory-Specific latent factor, which explained additional variance in working memory span tasks not captured by Common EF ($r = .98$, genetic correlation = 1.0, nonshared environmental correlation = .88). There was a large mean-level performance decline in Common EF ($d = -.60$) but not Working Memory-Specific ($d = -.$

Correspondence concerning this article should be addressed to Daniel Gustavson, Department of Psychiatry and Center for Behavior Genetics of Aging, University of California San Diego, 9500 Gilman Dr. (MC 0738), La Jolla, CA 92093. dgustavson@ucsd.edu. 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; Jeremy A. Elman, Department of Psychiatry, Center for Behavior Genetics of Aging, University of California, San Diego; 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; Kristen C. Jacobson, Department of Psychiatry and Behavioral Neuroscience, University of Chicago; Naomi P. Friedman, Department of Psychology and Neuroscience, Institute for Behavior Genetics, University of Colorado Boulder; Hong Xian, Department of Biostatistics, St. Louis University and Clinical Epidemiology Center, Veterans Affairs St. Louis Healthcare System; Rosemary Toomey, Department of Psychological and Brain Sciences, Boston University; 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 Center of Excellence for Stress and Mental Health, Veterans Affairs San Diego Healthcare System.

03). These results suggest that there is substantial decline in Common EF abilities across middle age, but that individual differences are almost perfectly stable.

Keywords

executive control; cognitive control; heritability; twin study; longitudinal design

Executive functions (EFs) are important cognitive abilities that control and regulate behavior (Friedman & Miyake, 2017; Miyake & Friedman, 2012; Miyake et al., 2000). Research on EFs in adolescence and young adulthood suggests that individual differences in EFs are highly stable over time at both the phenotypic and genetic levels (Friedman et al., 2016; Friedman, Miyake, Robinson, & Hewitt, 2011). This is especially true for those general EF abilities that underlie performance across a wide variety of EF tasks (Common EF). Our work has suggested that Common EF continues to underlie variation across a wide range of EF tasks in midlife (mean age 56 years), and is similarly correlated with other cognitive abilities as it is in studies of young adults (Gustavson et al., 2017a). Moreover, almost half of the variance of Common EF in midlife is still explained by genetic influences (heritability, or $a^2 = .46$), but environmental influences may account for a larger portion of the variation in Common EF in midlife compared to earlier ages (Gustavson et al., 2017a). However, it remains unclear whether individual differences in Common EF remain strongly correlated throughout middle age as in the first few decades of life, or whether new genetic and/or environmental influences arise as some individuals begin to experience age-related decline in EF.

In this longitudinal study, we examine the phenotypic and genetic/environmental stability of EFs in 1464 male-male twins who completed a neuropsychological battery that included seven EF tasks at up to two separate time-points (mean age 56 and 62 years) as part of the Vietnam Era Twin Study of Aging (VETSA). We expected that individual differences in Common EF would be highly correlated between these ages, especially at the genetic level. We also expected similar genetic stability of individual differences in Working Memory (WM)-Specific, a latent factor that accounts for EF abilities unique to working memory span above and beyond Common EF. Finally, we examined evidence for mean-level changes in the latent EF factors. We expected that, like other cognitive abilities, EFs would decline over this six-year interval even if individual differences remain stable.

The Unity and Diversity Model of Executive Function and its Stability in Early Life

Current theoretical conceptions of individual differences in EFs highlight two broad types of EF processes (Friedman & Miyake, 2017; Miyake & Friedman, 2012). One set of processes, called Common EF, are general abilities that underlie variation across a wide range of individual EFs including prepotent response inhibition (inhibition), task-set shifting (shifting), WM span, and WM updating (Engelhardt, Briley, Mann, Harden, & Tucker-Drob, 2015; Friedman et al., 2016; Friedman et al., 2008; Gustavson et al., 2017a). These Common EF abilities are thought to support general goal management, maintenance, and

implementation (Friedman & Miyake, 2017; Miyake & Friedman, 2012), and may be the aspects of EF that are most relevant to other phenomena such as psychopathology (Gustavson et al., 2017b; Herd et al., 2014; Snyder, Miyake, & Hankin, 2015), everyday self-regulation and goal management (Gustavson, Miyake, Hewitt, & Friedman, 2015), and even the expression of implicit racial biases (Ito et al., 2015).

There are also EF-specific processes that explain variation in individual EF abilities above and beyond Common EF. For example, Shifting-Specific processes account for additional variation in set-shifting tasks not captured by Common EF; it is thought to reflect the speed with which one can replace goals in WM when new ones become relevant (Friedman & Miyake, 2017; Miyake & Friedman, 2012). Similarly, Updating-Specific processes account for variation in WM updating tasks above and beyond Common EF, and are thought to reflect the effective gating of information into WM by the basal ganglia (Friedman & Miyake, 2017; Miyake & Friedman, 2012). In summary, EFs in a wide range of domains show substantial unity (i.e., Common EF underlying multiple EF processes), but also have considerable diversity (e.g., Updating-Specific and Shifting-Specific).

Most of the research on EFs using this so-called unity and diversity framework has examined Common EF and specific EF processes at the latent construct level in genetically informative samples using a battery of 7–12 EF tasks (Engelhardt et al., 2015; Friedman et al., 2016; Friedman et al., 2008; Gustavson et al., 2017a). In adolescence, individual differences in Common EF were almost exclusively explained by genetic influences ($a^2 = .96$ to 1.0; Engelhardt et al., 2015; Friedman et al., 2016). There were no significant shared environmental influences, which make twins more similar to one another ($c^2 = .00$), nor nonshared environmental influences, which make twins different ($e^2 = .00 - .04$). In early adulthood (mean age 23 years), Common EF was again explained mostly by genetic influences ($a^2 = .81$). At this age, there was some variance captured by shared environmental factors ($c^2 = .04$) and nonshared environmental factors ($e^2 = .15$), but only the latter were significant (Friedman et al., 2016).

In addition to showing that almost all the variance in Common EF is captured by genetic influences in the first few decades of life, this work has also revealed that these genetic influences are highly stable over time (Friedman et al., 2016; Friedman et al., 2011). For example, the genetic influences on Common EF in late adolescence (mean age 17 years) and early adulthood (mean age 23 years) were perfectly correlated with one another (genetic correlation, or $r_g = 1.0$, phenotypic correlation $r = .86$), suggesting that Common EF is explained by the same genetic influences in adolescence and adulthood (Friedman et al., 2016). In early childhood between ages 14 and 36 months, twins in this sample with greater self-restraint (i.e., longer latency to touch an attractive toy when prohibited) had significantly greater Common EF in adolescence (total $N = 945$; Friedman et al., 2011). This longitudinal relationship was due to a significant genetic correlation of early self-restraint with age 17 Common EF ($r_g = .49$) rather than a nonshared environmental correlation ($r_e = .21$), and Common EF mediated the association between childhood self-restraint and intelligence in adolescence. Based on these existing studies, it seems that Common EF shows considerable stability in early stages of life, especially at the genetic level, even as

mean-level performance on EF tasks continues to improve into young adulthood (Friedman et al., 2016).

Executive Functions in Midlife

Although there is evidence for the stability of Common EF in early life, little research has examined the stability of the genetic and environmental influences of EFs in later stages of life. Inhibition, shifting, and WM updating tasks continue to show strong phenotypic overlap in middle age and late life (ages 60 – 90; Vaughan & Giovanello, 2010), and a single general factor may underlie the phenotypic associations between these EFs in mid-late adulthood (ages 53 – 90; de Frias, Dixon, & Strauss, 2006, 2009). Our recently published work was the first examination of the genetic/environmental etiology of a latent Common EF factor in middle age (ages 51 – 61; Gustavson et al., 2017a). This model is displayed in Figure 1.

As shown in Figure 1, a Common EF factor explained variation in all seven neuropsychological tasks, and a WM-Specific factor explained additional variation in three WM span tasks above and beyond Common EF.[1] There was no evidence for EF-specific processes unique to inhibition or shifting. The lack of an Inhibition-Specific factor was consistent with research in younger samples (Friedman et al., 2016), but it is possible that a Shifting-Specific factor was not observed because the neuropsychological measures of shifting differ from the rapidly-paced task-switching paradigms used in younger samples (Friedman et al., 2016).

Also depicted in Figure 1, we observed that almost half of the variation in Common EF could be explained by genetic influences ($a^2 = .46$). These genetic influences were moderately correlated with genetic influences on general cognitive ability assessed in both midlife (age 56, $r_g = .59$) and young adulthood (age 20, $r_g = .45$). Because genetic influences on general cognitive ability are perfectly stable over this 35-year interval ($r_g = 1.0$; Lyons et al., 2017; Lyons et al., 2009), these results provide some initial evidence that the genetic influences on Common EF in middle age are similar to those identified in young adulthood, at least with regard to their overlap with other cognitive abilities.

There were also significant shared environmental influences ($c^2 = .41$) and nonshared environmental influences ($e^2 = .13$) on Common EF in midlife, the former of which had not been previously observed at earlier ages. Both types of environmental influences were also strongly correlated with the environmental influences on general cognitive ability measured in early adulthood ($r_c = .99$, $r_e = .72$). These environmental influences explained a relatively small portion of variance in general cognitive ability at either wave, especially for shared environment ($c^2 = .14$ to .16; $e^2 = .26$ to .29). Nevertheless, these results suggest that environmental influences demonstrate some stability throughout adulthood.

Although these results provide some evidence for the stability of genetic/environmental influences on Common EF across adulthood, it remains unclear whether individual

---

[1]In contrast to the studies of children and adolescents, we refer to a WM-Specific factor rather than an Updating-Specific factor because our WM tasks did not have a strong updating component, though WM span and WM updating seem to be highly similar (Schmiedek, Hildebrandt, Lövdén, Wilhelm, & Lindenberger, 2009).

differences in Common EF continue to show similar stability throughout middle age. Specifically, mid-to-late life is marked by a steady decline in cognitive abilities (Harris & Deary, 2011; Salthouse, 2005), and to the extent that some individuals decline earlier or more rapidly than others, this would suggest that new genetic and/or environmental influences may influence EFs during middle age. However, if most individuals decline to the same extent, there could be few if any new genetic/environmental influences on EFs even as individuals age. This latter possibility is most consistent with research in adolescence and young adults (i.e., mean-level developmental changes but little to no new genetic/ environmental influences), though it is possible age-related decline in EF will follow a different trajectory than its development.

Few studies have quantified age-related decline in EF other than its intersection with WM (Hertzog, Dixon, Hultsch, & MacDonald, 2003; Salthouse, 2005; Salthouse, Atkinson, & Berish, 2003). In a longitudinal analysis of the Victoria Longitudinal Study, for example, Hertzog et al. (2003) found evidence for change in WM abilities across a 6-year interval in individuals ranging from 61 to 91 years. At the first assessment, their WM latent variable was negatively correlated with age ($r = -.19$). A latent variable comprising the 6-year change in WM was also negatively associated with age ($r = -.27$), and much of this variance in WM change could be accounted for by age-related change in general cognitive ability.

In another investigation of the Victoria Study, a latent EF variable comprising two inhibition and two shifting tasks was negatively associated with age ($\beta = -.48$, controlling for vocabulary and fluid intelligence; de Frias et al., 2006), suggesting a steady decline in Common EF between middle age and old adulthood (aged 55 – 85). Further research suggested that this Common EF factor also accounts for variation in two WM updating tasks (de Frias et al., 2009) in cognitively normal, cognitively impaired (i.e., performing worse than 1.5 *SD* below the mean in at least one of five other cognitive domains), and cognitively elite individuals (i.e., performing above the mean in all five domains). For cognitively normal and cognitively elite individuals, a three-factor model of inhibition, shifting, and updating fit better than a unitary Common EF factor, which fit best for impaired individuals. However, this study did not examine models with Common EF and EF-specific variance components directly (e.g., WM-Specific), only whether there was one unitary factor or three correlated factors. Importantly, at a 3-year follow-up, there was evidence for longitudinal invariance of their best-fitting model in all three groups, suggesting that the EF factor structure is stable at least in short-term intervals in older age (de Frias et al., 2009).

These previous studies compared factor structures over time in older adults and correlations with age. However, it will also be useful to quantify the mean-level decline in latent EF factors across waves of assessment and examine the extent to which the genetic/ environmental components of their individual differences demonstrate stability or change. It will also be important to examine these associations using samples with narrow age ranges, especially in midlife when it is unclear how rapidly these abilities are declining. Finally, examining these associations in the context of the unity and diversity model will be useful in directly examine change in specific abilities (i.e., WM-Specific).

The Current Study

The current study examined the stability of the genetic/environmental influences and the mean-level changes in Common EF and WM-Specific factors across a six-year interval in middle age (mean age 56–62). We expected that both the genetic and environmental influences on Common EF and WM-Specific would be highly stable over the six-year period. Such findings would be consistent with the phenotypic and genetic stability of EFs at earlier ages (Friedman et al., 2016) as well as the near perfect genetic stability of general cognitive abilities in this age range (Lyons et al., 2017; Tucker-Drob & Briley, 2014), which share some overlapping genetic/environmental variance with EFs (Gustavson et al., 2017a). However, to the extent that we observed new variance in EFs, we expected that this new variance would be due to environmental influences rather than new genetic influences, as nonshared environmental influences accounted for the majority of change in Common EF between adolescence and adulthood (Friedman et al., 2016), as well as for cognitive abilities more generally across middle age (Kremen, Moore, Franz, Panizzon, & Lyons, 2014; Lyons et al., 2017; Tucker-Drob & Briley, 2014).

We also expected to observe mean-level decline in EF. These findings would be consistent with correlational evidence from WM and EFs across middle and old ages and the fact that other cognitive processes begin to steadily decline as early as the 50s (Rönnlund & Nilsson, 2006; Salthouse, 2005; Salthouse et al., 2003). This is also expected given theoretical proposals and empirical evidence that prefrontal cortical regions associated with EFs are some of the first and most strongly affected in normal cognitive aging (Bakkour, Morris, Wolk, & Dickerson, 2013; Buckner, 2004; Fjell et al., 2009). Although this work has not been integrated into the unity and diversity model of EFs, it is most likely that Common EF reflects these prefrontal cortical processes that are most sensitive to aging and likely declining by middle age. We did not make specific predictions regarding decline in WM-Specific as the decline in WM noted in previous studies may be driven by variance in Common EF, WM-Specific, or some combination of both (Hertzog et al., 2003).

Method

Subjects

Analyses were based on 1464 male twins (851 monozygotic [MZ] and 613 dizygotic [DZ] twins) from the longitudinal VETSA project. Twins were included in these analyses if they completed at least one of the two waves of assessment, either at wave 1 ($N = 1285$, M = 55.89 years, $SD = 2.44$ years), or wave 2 ($N = 1193$, $M = 61.73$, $SD = 2.44$), though most subjects completed both waves of assessment ($N = 1014$). VETSA participants all served in the United States military at some point between 1965 and 1975, and were recruited through random selection from the Vietnam Era Twin Registry from a previous study (Tsuang, Bar, Harley, & Lyons, 2001). Individuals in the VETSA are generally representative of the population of American men in their age group with respect to health and lifestyle factors, and nearly 80% of the sample did not serve in combat or in Vietnam (Kremen et al., 2011; Kremen et al., 2006; Schoenborn & Heyman, 2009). All data collection was approved by Institutional Review Boards at the participating institutions.

Measures

All of the dependent measures were adjusted for age by creating residualized scores after accounting for the effect of age (Gustavson et al., 2017a). Additionally, all dependent measures were standardized based on scores at the first wave of assessment. Therefore, means and variances at the second wave of assessment reflect the change in performance on the EF tasks across the six-year interval (see Table 1 for descriptive statistics).

For the second wave of assessment, we also adjusted scores to account for practice effects according to the method of Rönnlund, Nyberg, Bäckman, and Nilsson (2005). Practice effects were computed for each task based on the difference between wave 1 and wave 2 scores for individuals who returned ($N = 1014$) compared to data from matched attrition-replacement subjects who took the tests for the first time at wave 2 ($N = 179$), while also accounting for attrition effects using data from individuals who did not return at wave 2 ($N = 271$).[2] By accounting for practice effects, we were not only able to adjust the scores of the follow-up subjects to account for their repeated exposure, but were also able to utilize the attrition-replacement subjects to help fit the model at the second wave without fear that they would bias the results. Attrition effects are also accounted for in the computation of the practice effect for each task.

Inhibition—Inhibition was assessed with two tasks: (a) the Golden and Freshwater (2002) verson of the Stroop task (Stroop, 1935) and (b) 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 words identified during the color-word condition after adjusting for performance on the word-only and color-only conditions. The dependent measure of the AX-CPT was an arcsine-transformed signal detection index (d') based on the hit rate for *AX* trials minus the false alarm rate for *BX* trials. Additionally, we trimmed all d' prime values less than 0 to 0 to reduce the tail of the distribution (Gustavson et al., 2017a).

Working memory span—WM span was assessed with three tasks: (a) the letter-number sequencing subtest of the Wechsler Memory Scale-III (Wechsler, 1997), (b) the reading span task (Daneman & Carpenter, 1980) and (c) the forward and backward digit span subtests of the Wechsler Memory Scale-III (Wechsler, 1997). For letter-number sequencing and digit span, the dependent measure was the total number of trials passed. For the reading span, the dependent measure was the total number of correct words recalled across the entire task (5 trials each of length 2, 3, and 4 sentences).

Shifting—Shifting was assessed using two tasks from the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001): (a) the Trail Making Test and (b) the category switching trial of the verbal fluency test. For the Trail Making Test, the dependent measure was the time taken to complete the switching trial (trial 4) after residualizing the time on the single-task trials (trials 2 and 3, number sequencing and letter sequencing, respectively). This measure was reverse scored in all analyses so that higher numbers indicated better performance (like all other EF tasks). For category switching, the dependent

---

[2]Practice effects were computed for all EF tasks, though they were only significant for letter-number sequencing and digit span (see Table 1).

measure was a residualized score for category switching accuracy (the number of times a participant correctly switched categories) after adjusting for the number of correct responses across both category fluency trials (animals and boys' names).

Data Analysis

All analyses were conducted using the structural equation modeling package OpenMX in R (Boker et al., 2011), which accounts for missing observations using full-information maximum likelihood. Model fit was determined using the −2 log-likelihood values (−2LL), Bayesian information criterion (BIC), and the Root Mean Square Error of Approximation (RMSEA). Good fitting models were determined based on the lowest values for the BIC, and RMSEA values less than .06 (Hu & Bentler, 1998; Markon & Krueger, 2004). In addition to these statistics, multivariate models were also compared to the full genetic Cholesky decompositions (using χ^2 difference tests) to show that they did not fit worse than these full Cholesky models. Significance of individual parameters was established using likelihood-based 95% confidence intervals (95% CI) or with χ^2 difference tests (by fixing those parameters to zero).

Genetic analyses were based on the following classical assumptions in twin designs. Additive genetic influences (A) are assumed to correlate at 1.0 for MZ twins and at 0.5 for DZ twins because MZ twins share 100% of their alleles identical-by-descent and DZ twins share, on average, 50% of their segregating alleles identical-by-descent. Shared environmental influences (C), which make twins more similar, are assumed to correlate at 1.0 for both types of twins. Nonshared environmental influences (E), which make twins dissimilar (and also include measurement error for non-latent variables), are not correlated in either MZ or DZ twins by definition. We also assume that means and variances are identical across twin pair (twin 1 vs. twin 2) and across zygosity (MZ vs. DZ twins). These standard assumptions for univariate twin models extend to the multivariate models described here. In the longitudinal models, the phenotypic correlations between the latent factors are decomposed into genetic (r_g), shared environmental (r_c), and nonshared environmental correlations (r_e) by fitting a Cholesky decomposition.

In the confirmatory models presented here, we examined the longitudinal stability of EFs in a similar way to Friedman et al. (2016). Before fitting the longitudinal model of EFs across waves of assessment, we first confirmed that the common pathway model from the first wave continued to provide an adequate fit to the data at the second wave. Next, we combined the models at both waves. In this longitudinal model, the Common EF and WM-Specific latent factors, which are orthogonal within-wave, were also constrained to be orthogonal across wave. Furthermore, it is necessary to estimate residual correlations between individual tasks (e.g., Stroop at wave 1 and Stroop at wave 2), to account for within-task correlations not captured by the latent factors (Friedman et al., 2016).

Additional analyses were conducted to examine the potential effect of extreme scores on the results. The longitudinal models presented in the results were also examined after removing observations for each EF task when a score was above or below 3 *SD* from the mean. Individual differences and mean-level change estimates were nearly identical to the estimates presented in the results, so we present the non-trimmed data here.

Results

Descriptive Statistics and Preliminary Analyses

Descriptive statistics for all the measures are summarized in Table 1. The full phenotypic correlation matrix between all measures is displayed in Table 2 (below the diagonal), alongside the Twin 1 – Twin 2 correlations for MZ and DZ pairs (above the diagonal). Phenotypic correlations between the same tasks at wave 1 and wave 2 were moderate-to-strong (median $r = .57$), suggesting that there was considerable stability of individual differences at the task level. As shown in Table 1, all means at the second wave of assessment were significantly lower than 0 ($ts > 4.15$, $ps < .001$), indicating that performance declined on all EF tasks over the six-year interval.

The common pathway model at wave 2 alone is displayed in the supplemental material (Figure S1) alongside model comparisons for alternative models (Table S1). In summary, the unity and diversity model of EF at the second wave of assessment fit the data well, and was similar to the model from the first wave. There was a Common EF and a WM-Specific factor, with no evidence for an Inhibition-Specific or a Shifting-Specific factor. Because the results of this model are contained within those of the full longitudinal model, we do not discuss this model further.[3]

Longitudinal Model of Executive Function

The longitudinal model of Common EF and WM-Specific is displayed in Figure 2. This model had acceptable fit, −2LL = 40835.83, df = 16533, BIC = −68785, RMSEA = .011, and did not fit worse than the multivariate Cholesky decomposition, $\chi^2_{diff}(214) = 137.96$, $p = .999$. An alternate version of this model with unstandardized factor loadings is also presented in the supplement (Figure S2).

Individual differences—Consistent with our expectations, individual differences in Common EF were highly stable over time. The estimated phenotypic correlation between Common EF at age 56 and age 62 was $r = .97$, 95% CI [.93, .99]. Genetic influences on Common EF were perfectly correlated across time, $r_g = 1.0$, 95% CI [.89, 1.0], and accounted for almost half of the variation in Common EF at both waves of assessment, $a^2 = .42$, 95% CI [.15, .74] for wave 1; $a^2 = .46$, 95% CI [.15, .71] for wave 2. Similarly, shared environmental influences were also perfectly correlated over time, $r_c = 1.0$, 95% CI [.89, 1.0]. These shared environmental influences explained a similar portion of variance as the genetic influences at both waves, $c^2 = .45$, 95% CI [.21, .68] for wave 1; $c^2 = .27$, 95% CI [.06, .52] for wave 2. Nonshared environmental influences were also highly correlated over time, $r_e = .95$, 95% CI = [.75, 1.0], though they explained a significantly larger portion of the

---

[3]There are a few noticeable differences in the results of the model displayed in Figure S1 and the full longitudinal model from Figure 2. Namely, the estimates differ on WM-Specific, especially for the genetic ($a^2 = .04$, 95% CI = .00, .84) and shared environmental influences ($c^2 = .67$, 95% CI = .00, .90), which have wide confidence intervals. However, the estimates described in Figure 2 are within the confidence intervals of these estimates. The differences in these estimates are attributable to the low power to detect differences between genetic and shared environmental influences, even in this large sample (Martin, Eaves, Kearsey, & Davies, 1978). Including data from wave 1 provides additional information (i.e., the cross-wave cross-twin correlations) that helps discriminate these influences. For this reason, we focus our interpretations on these estimates for the longitudinal model.

variation at the second wave compared to the first wave, $e^2 = .13$, 95% CI [.06, .22] for wave 1; $e^2 = .26$, 95% CI [.17, .37] for wave 2, $\chi^2_{diff}(1) = 5.68$, $p = .017$.

Individual differences in WM-Specific were also highly stable over time, $r = .98$, 95% CI [.93, 1.0]. Genetic influences on WM-Specific were perfectly correlated over time, $r_g = 1.0$, 95% CI [.98, 1.0]. These genetic influences accounted for about three-quarters of the variation at both waves, $a^2 = .77$ [95% CI = .45, .92] for wave 1; $a^2 = .76$, 95% CI = [.39, .94] for wave 2. The shared environmental influences were also perfectly correlated over time, $r_c = 1.0$, 95% CI [−.06, 1.0], though this correlation was nonsignificant because the shared environmental influences accounted for only a small and nonsignificant portion of variance in WM-Specific at either wave, $c^2 = .06$, 95% CI [.00, .34] for wave 1; $c^2 = .08$, 95% CI [.00, .40] for wave 2. Nonshared environmental influences were also strongly correlated over time, $r_e = .88$, 95% CI [.46, 1.0]. They accounted for 16% of the variation in WM-Specific at both waves, $e^2 = .16$, 95% CI [.07, .25] for wave 1; $e^2 = .16$, 95% CI [.03, .29] for wave 2.

In total, genetic influences accounted for 46% of the phenotypic stability on Common EF (i.e., .44 of the .97 phenotypic correlation), with shared and nonshared environmental influences accounting for 36% and 18% of the phenotypic stability, respectively. For WM-Specific, genetic influences accounted for 78% of the phenotypic stability. Shared and nonshared environmental influences accounted for 7% and 14% of the phenotypic stability, respectively. As noted by the overlapping confidence intervals, genetic and environment influences generally explained the same proportion of variance in both EF factors over time. In fact, the six genetic/environmental variance components could be equated simultaneously without a significant reduction in fit, $\chi^2_{diff}(6) = 10.36$, $p = .110$. However, we do not display this further constrained model because the nonshared environmental influences on Common EF explained a significantly larger portion of the variance at the second wave when compared individually, $\chi^2_{diff}(1) = 5.68$, $p = .017$.

The residual genetic and environmental influences on the seven tasks are displayed in Table 3, alongside the genetic/environmental correlations between these residuals for the same tasks over time. Most of the residual variances on the individual tasks were explained by nonshared environmental influences (which include measurement error). Residual phenotypic correlations between EF tasks at wave 1 and wave 2 were small to moderate ($r = .08$ to .46) and were explained relatively equally by genetic and nonshared environmental influences, though for the most part only the nonshared environmental correlations were significant.

Mean-level differences—The model described in Figure 2 demonstrated configural invariance (i.e., the same latent constructs accounted for performance across EFs over time). To examine mean differences at the level of latent variables, it was necessary to impose some additional constraints on the model regarding factorial invariance. These analyses are described in Table 4, and followed procedures described by Widaman, Ferrer, and Conger (2010). Specifically, we equated the unstandardized factor loadings for each of the EF tasks over time (weak factorial invariance) and then further constrained the intercepts of the EF tasks to be equal over time (strong invariance). This strong invariance model is displayed in

Figure 3 (and Model 4 of Table 4), and an alternate depiction with unstandardized factor loadings is presented in the supplement (Figure S3). Although this model had a significantly worse fit to the data than the model in Figure 2, $\chi^2_{diff}(13) = 85.89$, $p < .001$, it had an equivalent BIC value, and revealed qualitatively similar factor loadings and individual differences results as the model displayed in Figure 2. Therefore, the significant drop in fit was likely due to the large sample size and high power to detect small deviations in observed versus predicted correlations and means (West, Taylor, & Wu, 2012), rather than a poor fit of the model (for a similar example, see Tucker-Drob, Briley, Starr, & Deary, 2014).

The results displayed in Figure 3 suggest that the mean changes across the EF tasks were due to a decrease in Common EF, rather than WM-Specific. The means and variances of both latent variables are standardized at the first wave of assessment. Therefore, the mean of Common EF ($-.60$) at the second wave can be interpreted as a decrease in .60 *SD* compared to the first wave, 95% CI [$-.68$, $-.53$]. The variance of Common EF did not change at the second wave of assessment, 1.04, 95% CI [.90, 1.15]. In contrast to these results for Common EF, there was little evidence for mean changes in WM-Specific, $-.03$, 95% CI [$-.13$, .07]. However, the variance of WM-Specific was smaller at the second wave compared to the first wave, .77, 95% CI [.65, .88].

Discussion

The current study was the first to examine the phenotypic, genetic, and environmental stabilities and mean-level decline in Common EF and WM-Specific abilities in middle age. We showed that individual differences in Common EF and WM-Specific remained highly stable over the six-year window, especially the genetic and shared environmental correlations, which were estimated at 1.0. Additionally, the mean-level decline in performance across EF tasks was due to a drop in Common EF but not WM-Specific. These results suggest that, despite a decline in EF ability over time, individual differences in EF exhibit remarkable stability.

Implications for the Stability of Individual Differences in Executive Functions in Midlife

Common EF—Prior research indicates that individual differences in Common EF show strong stability between adolescence and young adulthood ($r = .97$), especially with regard to genetic influences, which are perfectly correlated across time ($r_g = 1.0$; Friedman et al., 2016). The results of the current study extend these previous findings by showing that genetic influences on Common EF continue to show high stability across a six-year interval in middle age. Therefore, as yet, there is no evidence for new genetic influences on Common EF as individuals age. This parallels our findings regarding genetic influences in general cognitive ability over time (Lyons et al., 2017; Lyons et al., 2009). Moreover, these results suggest that ongoing gene-discovery efforts regarding EFs should consider models akin to Common EF that capture stable genetic variance across multiple EF situations. These factors should provide a better phenotype than individual EF tasks, which include domain-specific variance (WM-Specific), task-specific variance (e.g., unique to the Stroop), and measurement error.

Additionally, these findings provide further evidence for the newly identified shared environmental influences on Common EF. Earlier work has suggested that there are no shared environmental influences on Common EF in adolescence (Engelhardt et al., 2015; Friedman et al., 2008), and only weak and nonsignificant evidence for shared environmental influences on Common EF in young adulthood ($c^2 = .04$; Friedman et al., 2016). Here, we observed significant shared environmental influences at both waves. Our previous work showed that these shared environmental influences were significantly correlated with those on early adult general cognitive ability ($r_c = .99$; Gustavson et al., 2017a), suggesting they are not new to middle age. This finding remains somewhat puzzling because we would expect shared environmental influences to weaken as the childhood familial environment becomes more distal. In contrast, these shared environmental influences explained a relatively weak proportion of variance in general cognitive ability in young adulthood (14%), but a substantial portion of variance in Common EF at mean ages 56 (46%) and 62 (27%), though the confidence intervals were somewhat broad. Nevertheless, these results extend these findings by suggesting that they also remain perfectly stable, at least over the six-year time frame tested here.

These results also provided some evidence for the stability of nonshared environmental influences on Common EF. Research by Friedman et al. (2016) revealed that the nonshared environmental influences on Common EF were moderately correlated between late adolescence and early adulthood ($r_e = .39$), but this correlation was not significant. In the current study, environmental influences were strongly correlated with one another ($r_e = .95$), but they explained a significantly larger portion of the variation at the second wave (26%) compared to the first wave (13%). Therefore, the nonshared environmental influences on EF may demonstrate higher stability in later stages of adulthood than in adolescence or early adulthood. However, these similar environmental influences appear to exert a stronger influence on Common EF later in middle age than early adulthood or early middle age, although this could be also due to a weaker influence of genetic or shared environmental factors.

WM-Specific—In general, the conclusions regarding WM-Specific were nearly identical to those for Common EF. The overall phenotypic correlation between the WM-Specific factors between waves of assessment was nearly perfect ($r = .98$). Like Common EF, this phenotypic stability was due to identical genetic and shared environmental influences, and a strong correlation for the nonshared environmental influences ($r_e = .88$).

This is the first study to model longitudinal change in WM-Specific processes at the latent variable level, so it is somewhat unclear how these findings map onto those from adolescence and young adulthood. However, previous work has examined the stability of WM Updating-Specific processes using complex WM tasks that do not focus as solely on span (Friedman et al., 2016; Friedman et al., 2008), but which are empirically highly similar (Schmiedek et al., 2009). This previous work has suggested that the genetic influences on WM Updating-Specific are also highly conserved between late adolescence and early adulthood ($r_g = .99$; Friedman et al., 2016), so it is not surprising that we also observed perfect genetic stability on our WM-Specific factor in midlife.

Theoretical Implications for Mean-Level Decline in Executive Function in Midlife

These findings provide insights into the overall decline in cognitive abilities in aging. To our knowledge, this is the first study to quantify the mean-level changes in Common EF and WM-Specific in middle age using a longitudinal design. After accounting for practice effects there was considerable decline in Common EF, but not WM Specific over this six-year interval.

This substantial decline in Common EF may be somewhat surprising, but it is largely consistent with other estimates of age-related decline in EF and WM, and with theoretical perspectives that normal brain aging is especially pronounced in prefrontal cortical regions associated with EFs (Buckner, 2004; Fjell et al., 2009). Cross-sectional data from the Victoria Longitudinal Study suggested that Common EF may decline by as much as half a standard deviation every 8.5 years (the *SD* of age in their sample) between middle age and old adulthood (de Frias et al., 2006). The same group revealed a smaller correlation between WM span and age ($r = -.19$; Hertzog et al., 2003). This result is also consistent with our findings. Had we examined WM span factor alone, we would have likely observed a similar smaller total decline. By examining WM-Specific—which unlike WM Span, is independent of Common EF—our results suggest that decline in WM Span is probably due to its link with to Common EF. Finally, our results are consistent with cross-sectional associations within our sample. Reanalysis of our data without adjusting individual measures for age resulted in a negative association between age and Common EF at wave 1 ($r = -.13$) and wave 2 ($r = -.23$), mapping onto similar expected declines in Common EF over a six-year interval (i.e., $-.32$ and $-.57$ *SD*, respectively) and suggesting that the rate of decline was stronger in the early 60s than the late 50s.

Another reason that we observed a strong decline is that we accounted for practice effects. Reanalysis of the strong invariance model described in Figure 3 with data that were unadjusted for practice effects indicated that Common EF declined by only $-.20$ *SD* and WM-Specific by $-.08$ *SD*. Based on these findings, we conclude that not accounting for practice effects in longitudinal designs may result in substantial under-estimation of age-related declines in Common EF. Because practice effects are not often accounted for in cognitive aging research (Ferrer, Salthouse, McArdle, Stewart, & Schwartz, 2005; Nyberg, Lövdén, Riklund, Lindenberger, & Bäckman, 2012; Rönnlund & Nilsson, 2006; Rönnlund et al., 2005), especially in work on EFs and WM processes, future studies should further quantify the strength of the mean-level change in EF abilities and the extent to which practice effects mask this decline (e.g., familiarity with the stimuli and/or task requirements). The practice effect correction had little effect on the WM-Specific factor, suggesting that it is not as susceptible to practice effects at least at this age and across this follow-up interval. Further research will be needed to clarify what dimensions of WM processes are captured by the WM-Specific factor.

Some researchers have proposed that changes in certain cognitive abilities, such as processing speed, inhibition, or general cognitive ability may act as leading indicators of change in other cognitive domains (e.g., Hasher & Zacks, 1988; Salthouse, 1985; Tucker-Drob et al., 2014). Our results suggest that substantial age-related declines in Common EF begin at least as early as the mid-to-late 50s. Therefore, like processing speed, Common EF

may be predictive of age related declines in other cognitive abilities observed later in adulthood (e.g., Tucker-Drob et al., 2014). First, Common EF at age 56 was moderately correlated with processing speed (Gustavson et al., 2017a), but should be largely unconfounded with speed because dependent measures were adjusted for the baseline conditions that tap speed (e.g., Stroop, Trail Making Test). Nevertheless, it remains an open question whether the components of Common EF still associated with processing speed (even after controlling for speed within RT-based measures) are the aspects of Common EF that decline first.

Second, we observed no evidence for Inhibition-Specific abilities at either wave of assessment. Therefore, the extent to which decline in inhibition accounts for change in other cognitive domains (Hasher & Zacks, 1988) may be due to more general processes involved in Common EF rather than due solely to inhibition (Miyake & Friedman, 2012). Third, it is possible that the mean change in Common EF may reflect a decline in general cognitive ability (Tucker-Drob & Briley, 2014; Tucker-Drob et al., 2014). Common EF was moderately correlated with general cognitive ability at the first wave of assessment (r = .68, r_g = .59; Gustavson et al., 2017a), though this association may be larger in task batteries with greater proportions of WM span and WM updating tasks (Engelhardt et al., 2016). Interestingly, the mean decline in general cognitive ability in this sample was $d = -.32$ (using the Armed Forces Qualification Test reported in Gustavson et al., 2017), about half of the magnitude of the decline in Common EF. Our finding that Common EF declined more rapidly than general cognitive ability suggest that the mean level changes described here are not solely due to change in general cognition. However, it would be important to examine these possibilities using more informative models (e.g., dual change), especially those that can directly evaluate the direction of causation as it is possible that decline in EF precedes or underlies decline in general cognition rather than the reverse.

Combining both sets of results, the substantial mean-level decline but strong phenotypic, genetic, and environmental correlations for Common EF (rs = .95 to 1.0) suggest that there is likely no variability in change. In other words, individuals with greater Common EF at the first wave continued to have greater Common EF at the second wave even as their overall performance declined. Because there was also no change in total variance, it is unlikely that some parts of the distribution declined more rapidly than others, although this was possible for nonshared environmental influences which exhibited some change in variance explained at each wave. These findings are consistent with the strong stability of individual differences in Common EF throughout early life even as EFs improve into young adulthood (Friedman et al., 2016; Friedman et al., 2011), and suggest a complementary pattern in later adulthood: Individual differences remain highly stable even as they show considerable decline. These patterns of results are also in contrast with other cognitive abilities such as episodic memory, including results observed in the same subjects and time frame where there was considerable variability in cognitive decline at both the genetic and environmental levels (Panizzon et al., 2015).

These results should be interpreted in the context of some general limitations. First, it will be important to examine these associations in a sample that has females included. Second, we adjusted the second wave scores to account for practice effects (Rönnlund et al., 2005). We

acknowledge that our method may introduce some imprecision, but these calculations should have little effect on the individual differences results, and should be far better than excluding cases, controlling for dropout versus returner status, or ignoring practice effects altogether. Finally, there was some evidence that the strong invariance model (Figure 3), in which it was necessary to interpret mean-level change, fit significantly worse than the configural invariance model (Figure 2). Importantly, however, the BIC values were identical between both models, suggesting that they are equally good at balancing parsimony and fit. Moreover, the individual differences estimates (including genetic/environmental correlations) were nearly identical in both models, suggesting that interpreting this model did not impact the results.

Concluding Remarks

Even though neuropsychological measures of EF are widely used in research on cognitive aging, there is surprisingly little research on the stability of genetic/environmental influences on EFs in this age range (Kremen et al., 2014; Kremen et al., 2011). Our findings provide further evidence for the unity and diversity model of EFs in middle age, and suggest that genetic/environmental influences on individual differences in Common EF and WM-Specific remain highly stable during middle age. However, the strong stability of individual differences is contrasted with substantial mean-level decline, at least for Common EF. EFs are highly relevant to clinical and social outcomes (Gustavson et al., 2017b; Miyake & Friedman, 2012; Snyder et al., 2015), so the continued study of these constructs in relation to physical and mental health will be important in understanding the relevance of Common EF to cognitive aging.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Bakkour A, Morris JC, Wolk DA, Dickerson BC. The effects of aging and Alzheimer's disease on cerebral cortical anatomy: specificity and differential relationships with cognition. Neuroimage. 2013; 76:332–344. DOI: 10.1016/j.neuroimage.2013.02.059 [PubMed: 23507382]

Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, Fox J. OpenMx: An open source extended structural equation modeling framework. Psychometrika. 2011; 76:306–317. DOI: 10.1007/s11336-010-9200-6 [PubMed: 23258944]

Braver TS, Barch DM, Keys BA, Carter CS, Cohen JD, Kaye JA, Reed BR. Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. Journal of Experimental Psychology: General. 2001; 130:746–763. [PubMed: 11757878]

Buckner RL. Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate. Neuron. 2004; 44:195–208. doi: DOI: 10.1016/j.neuron. 2004.09.006 [PubMed: 15450170]

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior. 1980; 19:450–466.

de Frias CM, Dixon RA, Strauss E. Structure of four executive functioning tests in healthy older adults. Neuropsychology. 2006; 20:206–214. DOI: 10.1037/0894-4105.20.2.206 [PubMed: 16594781]

de Frias CM, Dixon RA, Strauss E. Characterizing executive functioning in older special populations: from cognitively elite to cognitively impaired. Neuropsychology. 2009; 23:778–791. DOI: 10.1037/a0016743 [PubMed: 19899836]

Delis, DC., Kaplan, E., Kramer, JH. Delis-Kaplan executive function system (D-KEFS). Psychological Corporation; 2001.

Engelhardt LE, Briley DA, Mann FD, Harden KP, Tucker-Drob EM. Genes unite executive functions in childhood. Psychological Science. 2015; 26:1151–1163. DOI: 10.1177/0956797615577209 [PubMed: 26246520]

Engelhardt LE, Mann FD, Briley DA, Church JA, Harden KP, Tucker-Drob EM. Strong genetic overlap between executive functions and intelligence. Journal of Experimental Psychology: General. 2016; 145:1141–1159. DOI: 10.1037/xge0000195 [PubMed: 27359131]

Ferrer E, Salthouse TA, McArdle JJ, Stewart WF, Schwartz BS. Multivariate modeling of age and retest in longitudinal studies of cognitive abilities. Psychology and Aging. 2005; 20:412–422. DOI: 10.1037/0882-7974.20.3.412 [PubMed: 16248701]

Fjell AM, Westlye LT, Amlien I, Espeseth T, Reinvang I, Raz N, Walhovd KB. High consistency of regional cortical thinning in aging across multiple samples. Cerebral cortex. 2009; 19:2001–2012. DOI: 10.1093/cercor/bhn232 [PubMed: 19150922]

Friedman NP, Miyake A. Unity and diversity of executive functions: Individual differences as a window on cognitive structure. Cortex. 2017; 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. Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. Developmental Psychology. 2016; 52:326–340. DOI: 10.1037/dev0000075 [PubMed: 26619323]

Friedman NP, Miyake A, Robinson JL, Hewitt JK. Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. Developmental Psychology. 2011; 47:1410–1430. DOI: 10.1037/a0023750 [PubMed: 21668099]

Friedman NP, Miyake A, Young SE, Defries JC, Corley RP, Hewitt JK. Individual differences in executive functions are almost entirely genetic in origin. Journal of Experimental Psychology: General. 2008; 137:201–225. DOI: 10.1037/0096-3445.137.2.201 [PubMed: 18473654]

Golden, CJ., Freshwater, SM. The Stroop Color and Word Test: A Manual for Clinical and Experimental Uses;[adult Version]. Stoelting; 2002.

Gustavson DE, Miyake A, Hewitt JK, Friedman NP. 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. 2015; 144:1063–1079. DOI: 10.1037/xge0000110 [PubMed: 26389573]

Gustavson DE, Panizzon MS, Franz CE, Friedman NP, Reynolds CA, Jacobson KC, Kremen WS. Genetic and environmental architecture of executive functions in midlife. Neuropsychology, Advance online publication. 2017a; doi: 10.1037/neu0000389

Gustavson DE, Stallings MC, Corley RP, Miyake A, Hewitt JK, Friedman NP. Executive functions and substance use: Relations in late adolescence and early adulthood. Journal of Abnormal Psychology. 2017b; 126:257–270. DOI: 10.1037/abn0000250 [PubMed: 28045282]

Harris SE, Deary IJ. The genetics of cognitive ability and cognitive ageing in healthy older people. Trends in Cognitive Sciences. 2011; 15:388–394. DOI: 10.1016/j.tics.2011.07.004 [PubMed: 21840749]

Hasher L, Zacks RT. Working memory, comprehension, and aging: A review and a new view. Psychology of Learning and Motivation. 1988; 22:193–225.

Herd SA, O'Reilly RC, Hazy TE, Chatham CH, Brant AM, Friedman NP. A neural network model of individual differences in task switching abilities. Neuropsychologia. 2014; 62:375–389. DOI: 10.1016/j.neuropsychologia.2014.04.014 [PubMed: 24791709]

Hertzog C, Dixon RA, Hultsch DF, MacDonald SW. Latent change models of adult cognition: Are changes in processing speed and working memory associated with changes in episodic memory? Psychology and Aging. 2003; 18:755–769. DOI: 10.1037/0882-7974.18.4.755 [PubMed: 14692862]

Hu LT, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods. 1998; 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. Toward a comprehensive understanding of executive cognitive function in implicit racial bias. Journal of Personality and Social Psychology. 2015; 108:187–218. DOI: 10.1037/a0038557 [PubMed: 25603372]

Kremen, WS., Moore, CS., Franz, CE., Panizzon, MS., Lyons, MJ. Cognition in middle adulthood. In: Finkel, D., Reynolds, CA., editors. Behavior genetics of cognition across the lifespan. Springer; New York: 2014. p. 105-134.

Kremen WS, Panizzon MS, Xian H, Barch DM, Franz CE, Grant MD, Lyons MJ. Genetic architecture of context processing in late middle age: More than one underlying mechanism. Psychology and Aging. 2011; 26:852–863. DOI: 10.1037/a0025098 [PubMed: 21875218]

Kremen WS, Thompson-Brenner H, Leung YM, Grant MD, Franz CE, Eisen SA, Lyons MJ. Genes, environment, and time: The Vietnam Era Twin Study of Aging (VETSA). Twin Research and Human Genetics. 2006; 9:1009–1022. DOI: 10.1375/183242706779462750 [PubMed: 17254445]

Lyons MJ, Panizzon MS, Liu W, McKenzie R, Bluestone NJ, Grant MD, Xian H. A longitudinal twin study of general cognitive ability over four decades. Developmental Psychology. 2017; 53:1170–1177. DOI: 10.1037/dev0000303 [PubMed: 28358535]

Lyons MJ, York TP, Franz CE, Grant MD, Eaves LJ, Jacobson KC, Kremen WS. Genes determine stability and the environment determines change in cognitive ability during 35 years of adulthood. Psychological Science. 2009; 20:1146–1152. DOI: 10.1111/j.1467-9280.2009.02425.x [PubMed: 19686293]

Markon KE, Krueger RF. An empirical comparison of information-theoretic selection criteria for multivariate behavior genetic models. Behavior Genetics. 2004; 34:593–610. DOI: 10.1007/s10519-004-5587-0 [PubMed: 15520516]

Martin NG, Eaves LJ, Kearsey MJ, Davies P. The power of the classical twin study. Heredity (Edinb). 1978; 40:97–116. DOI: 10.1038/hdy.1978.10 [PubMed: 272366]

Miyake A, Friedman NP. The nature and organization of individual differences in executive functions: Four general conclusions. Current Directions in Psychological Science. 2012; 21:8–14. DOI: 10.1177/0963721411429458 [PubMed: 22773897]

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive Psychology. 2000; 41:49–100. DOI: 10.1006/cogp.1999.0734 [PubMed: 10945922]

Nyberg L, Lövdén M, Riklund K, Lindenberger U, Bäckman L. Memory aging and brain maintenance. Trends in Cognitive Sciences. 2012; 16:292–305. DOI: 10.1016/j.tics.2012.04.005 [PubMed: 22542563]

Panizzon MS, Neale MC, Docherty AR, Franz CE, Jacobson KC, Toomey R, Kremen WS. Genetic and environmental architecture of changes in episodic memory from middle to late middle age. Psychology and Aging. 2015; 30:286–300. DOI: 10.1037/pag0000023 [PubMed: 25938244]

Rönnlund M, Nilsson LG. Adult life-span patterns in WAIS-R Block Design performance: Cross-sectional versus longitudinal age gradients and relations to demographic factors. Intelligence. 2006; 34:63–78. DOI: 10.1016/j.intell.2005.06.004

Rönnlund M, Nyberg L, Bäckman L, Nilsson LG. Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. Psychology and Aging. 2005; 20:3–18. DOI: 10.1037/0882-7974.20.1.3 [PubMed: 15769210]

Salthouse, TA. Speed of behavior and its implications for cognition. In: Birren, JE., Schaie, KW., editors. The handbooks of aging: Handbook of the psychology of aging. New York, NY: Van Nostrand Reinhold; 1985. p. 400-426.

Salthouse TA. Relations between cognitive abilities and measures of executive functioning. Neuropsychology. 2005; 19:532–545. DOI: 10.1037/0894-4105.19.4.532 [PubMed: 16060828]

Salthouse TA, Atkinson TM, Berish DE. Executive functioning as a potential mediator of age-related cognitive decline in normal adults. Journal of Experimental Psychology: General. 2003; 132:566–594. DOI: 10.1037/0096-3445.132.4.566 [PubMed: 14640849]

Schmiedek F, Hildebrandt A, Lövdén M, Wilhelm O, Lindenberger U. Complex span versus updating tasks of working memory: The gap is not that deep. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2009; 35:1089.

Schoenborn CA, Heyman KM. Health characteristics of adults aged 55 years and over: United States, 2004–2007. National Health Statistics Report. 2009; 16:1–31.

Snyder HR, Miyake A, Hankin BL. Advancing understanding of executive function impairments and psychopathology: Bridging the gap between clinical and cognitive approaches. Frontiers in Psychology. 2015; 6:328.doi: 10.3389/fpsyg.2015.00328 [PubMed: 25859234]

Stroop JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology. 1935; 18:643–662. DOI: 10.1037/0096-3445.121.1.15

Tsuang MT, Bar JL, Harley RM, Lyons MJ. The Harvard Twin Study of Substance Abuse: What we have learned. Harvard Review of Psychiatry. 2001; 9:267–279. DOI: 10.1093/hrp/9.6.267 [PubMed: 11600486]

Tucker-Drob EM, Briley DA. Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. Psychological Bulletin. 2014; 140:949–979. DOI: 10.1037/a0035893 [PubMed: 24611582]

Tucker-Drob EM, Briley DA, Starr JM, Deary IJ. Structure and correlates of cognitive aging in a narrow age cohort. Psychology and Aging. 2014; 29:236–249. DOI: 10.1037/a0036187 [PubMed: 24955992]

Vaughan L, Giovanello K. Executive function in daily life: Age-related influences of executive processes on instrumental activities of daily living. Psychology and Aging. 2010; 25:343–355. DOI: 10.1037/a0017729 [PubMed: 20545419]

Wechsler, D. WMS-III: Wechsler memory scale administration and scoring manual. San Antonio, TX: Psychological Corporation; 1997.

West, SG., Taylor, AB., Wu, W. Model fit and model selection in structural equation modeling. In: Hoyle, RH., editor. Handbook of structural equation modeling. 2012. p. 209-231.

Widaman KF, Ferrer E, Conger RD. Factorial Invariance Within Longitudinal Structural Equation Models: Measuring the Same Construct Across Time. Child Development Perspectives. 2010; 4:10–18. [PubMed: 20369028]
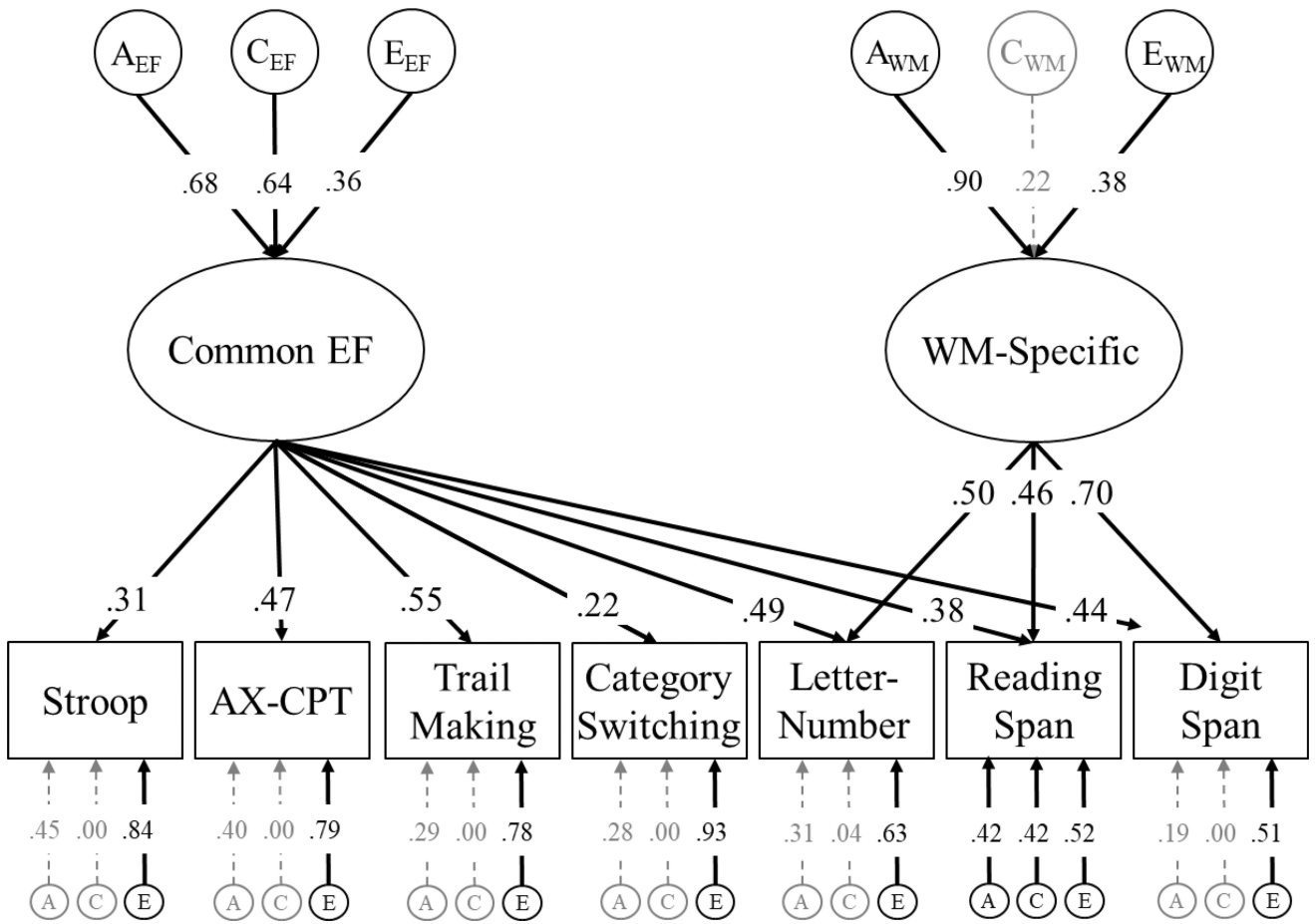
Figure 1.
Standardized estimates for genetic unity and diversity model of Common EF and WM-Specific abilities at the first wave of assessment (mean age 56), reproduced from (Gustavson et al., 2017a). The ACE factors represent genetic influences (A), shared environmental influences (C), and nonshared environmental influences (E) on each latent construct and individual measure. Ellipses indicate latent variables and rectangles indicate measured variables. Significant factor loadings are displayed in black and with solid lines ($p < .05$). Variation explained by latent factors can be computed by squaring the factor loadings.
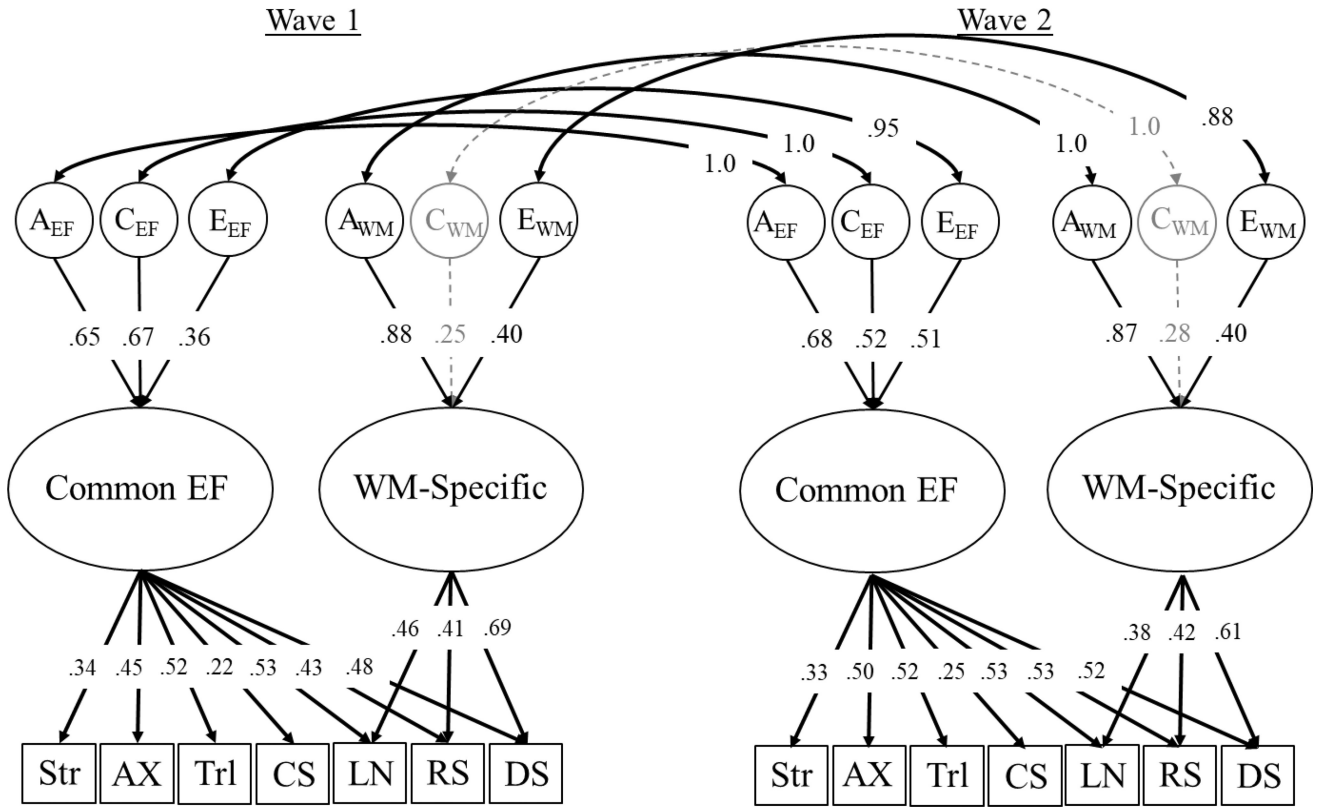
Figure 2.
Cross-wave model of the genetic (A), shared environmental (C) and nonshared environmental (E) influences on Common EF and WM-Specific latent variables. Ellipses represent latent variables and rectangles represent measured variables. Significant factor loadings and correlations are displayed with black arrows ($p < .05$). Residual ACEs are not displayed, but are instead presented in Table 3. All factor loadings are standardized, as are means and variances of all latent variables. Str = Stroop; AX = AX-Continuous Performance Test; Trl = Trail Making Test; CS = Category Switching Fluency; LN = Letter-number sequencing; RS = Reading Span; DS = Digit Span.
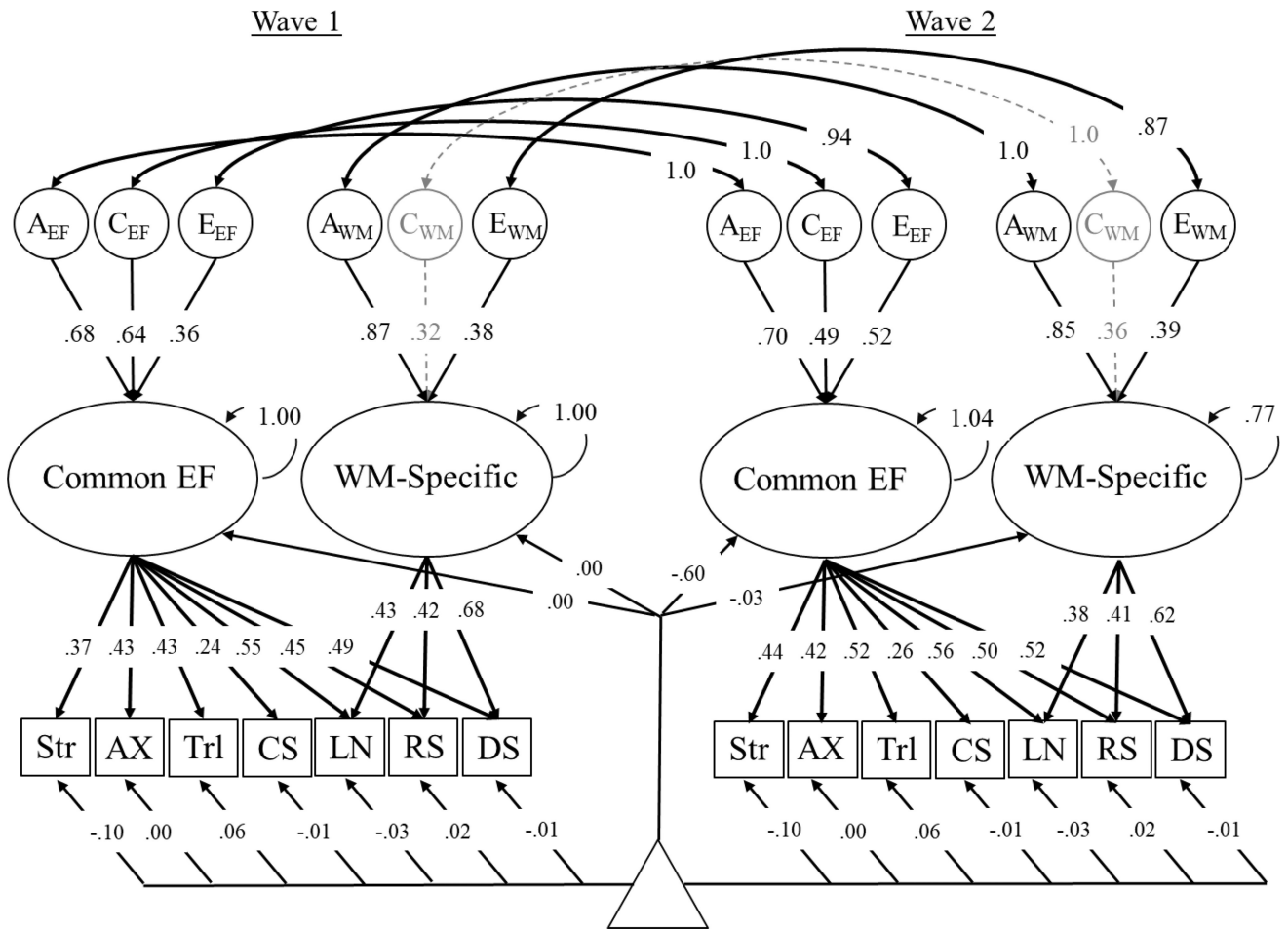
Figure 3.
Strong invariance model of the genetic (A), shared environmental (C) and nonshared environmental (E) influences on Common EF and WM-Specific latent variables. Ellipses represent latent variables, rectangles represent measured variables, and the triangle represents a constant used to model intercepts (i.e., capturing means). The circular arrows on Common EF and WM-Specific represent their unstandardized variances (wave 1 variances were fixed at 1.0 and wave 1 means were fixed at 0.0). In this strong invariance model, all unstandardized factor loadings and intercepts for EF tasks were equated over time, but factor loadings differ here because the standardized factor loadings are displayed. Significant factor loadings and correlations are displayed with solid black arrows ($p < .05$). Residual ACEs are not shown here, but were similar to those presented in Table 3 for the unrestricted model. Str = Stroop; AX = AX-Continuous Performance Test; Trl = Trail Making Test; CS = Category Switching Fluency; LN = Letter-number sequencing; RS = Reading Span; DS = Digit Span.

Table 1

Descriptive Statistics for All Measures in the Study

Task	N	M	SD	Range	Skewness	Kurtosis	Unstandardized M (SD)
Wave 1 (Age 56)							
Stroop	1250	0.00	1.00	−4.63 – 4.01	−0.12	0.63	35.95 (8.33)
AX-CPT	1190	0.00	1.00	−2.68 – 1.10	−1.35	1.06	.99 (.37)
Letter-Number	1280	0.00	1.00	−4.31 – 4.19	0.21	1.04	10.13 (2.36)
Reading Span	1248	0.00	1.00	−3.71 – 2.14	−0.53	0.05	34.14 (5.33)
Digit Span	1276	0.00	1.00	−2.38 – 3.35	0.29	−0.30	17.11 (3.90)
Trail-Making Test	1269	0.00	1.00	−4.98 – 2.56	−1.53	4.02	89.15 (35.05)
Category Switching	1275	0.00	1.00	−3.77 – 2.87	−0.44	0.58	11.52 (3.04)
Wave 2 (Age 62)							M change (no practice effect correction)
Stroop	1166	−0.37	0.85	−3.40 – 5.42	−0.20	2.36	−0.23
AX-CPT	1048	−0.13	1.04	−4.54 – 1.42	−1.42	2.96	0.10
Letter-Number	1188	−0.36	0.96	−3.35 – 3.23	0.11	0.47	−0.21
Reading Span	1183	−0.22	0.94	−3.35 – 2.04	−0.36	0.01	0.05
Digit Span	1047	−0.22	0.95	−2.54 – 3.25	0.35	−0.06	−0.09
Trail-Making Test	1036	−0.10	0.81	−4.09 – 2.33	−1.70	4.96	−0.10
Category Switching	1188	−0.17	0.97	−4.05 – 2.78	−0.56	1.00	−0.06

Note. All dependent measures are reported controlling for age and are standardized with respect to age. For wave 1, the final column shows the means and standard deviations before standardizing with respect to age. For wave 2, the final column shows the changes in means without accounting for practice effects (means are still shown standardized with respect to wave 1).

Table 2

Phenotypic and Cross-Twin Cross-Trait Correlations Between all Tasks

Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. W1 Stroop	**.29** / **.17**	.09 / .04	**.15** / **.13**	**.13** / .07	.11 / .09	**.13** / **.10**	.02 / .06	**.29** / **.17**	**.17** / .08	**.12** / .04	**.12** / .06	**.12** / .08	**.14** / **.12**	.05 / **.12**
2. W1 AX-CPT	**.16**	**.36** / **.18**	**.20** / **.11**	**.16** / **.13**	**.18** / **.16**	**.23** / **.19**	.10 / .06	.11 / −.01	**.31** / **.27**	**.16** / **.17**	**.14** / **.13**	**.16** / **.10**	**.15** / **.11**	.05 / .06
3. W1 Letter-Number	**.16**	**.23**	**.53** / **.36**	**.38** / **.25**	**.49** / **.34**	**.24** / **.26**	.07 / .06	.10 / .05	**.25** / **.22**	**.41** / **.26**	**.38** / **.28**	**.47** / **.28**	**.24** / **.24**	**.11** / .07
4. W1 Reading Span	**.16**	**.21**	**.43**	**.68** / **.49**	**.43** / **.28**	**.15** / **.16**	.07 / .04	**.14** / .09	**.22** / **.13**	**.33** / **.19**	**.52** / **.32**	**.42** / **.26**	**.16** / **.14**	**.11** / .05
5. W1 Digit Span	**.14**	**.18**	**.58**	**.50**	**.64** / **.39**	**.25** / **.23**	**.12** / .00	.10 / .03	**.21** / **.19**	**.41** / **.27**	**.45** / **.32**	**.62** / **.40**	**.23** / **.17**	**.10** / **.11**
6. W1 Trail-Making	**.16**	**.27**	**.28**	**.19**	**.26**	**.38** / **.21**	**.13** / −.03	**.14** / .07	**.17** / **.16**	**.21** / **.16**	**.19** / **.21**	**.21** / **.18**	**.28** / **.14**	.07 / **.10**
7. W1 Category Switch	.06	**.12**	**.12**	.09	**.10**	**.13**	**.14** / −.01	.02 / .00	**.10** / **.14**	.08 / .08	.09 / .04	**.13** / .01	.08 / −.02	**.16** / .08
8. W2 Stroop	**.50**	**.13**	**.15**	**.15**	**.14**	**.13**	.05	**.32** / **.14**	**.16** / .01	**.10** / .01	**.11** / .05	**.11** / .00	**.11** / .09	**.14** / **.12**
9. W2 AX-CPT	**.21**	**.57**	**.27**	**.26**	**.26**	**.26**	**.11**	**.15**	**.43** / **.35**	**.21** / **.17**	**.22** / **.11**	**.22** / **.17**	**.13** / **.13**	.06 / .09
10. W2 Letter-Number	**.16**	**.21**	**.58**	**.37**	**.51**	**.26**	**.11**	**.18**	**.26**	**.37** / **.19**	**.33** / **.22**	**.38** / **.26**	**.31** / **.12**	**.10** / **.11**
11. W2 Reading Span	**.18**	**.21**	**.46**	**.67**	**.55**	**.26**	**.11**	**.18**	**.28**	**.45**	**.57** / **.35**	**.41** / **.33**	**.21** / **.12**	**.11** / .08
12. W2 Digit Span	**.14**	**.21**	**.56**	**.46**	**.76**	**.27**	**.10**	**.17**	**.28**	**.51**	**.52**	**.57** / **.39**	**.27** / **.18**	**.13** / .02
13. W2 Trail-Making	**.19**	**.24**	**.27**	**.23**	**.24**	**.32**	**.10**	**.17**	**.24**	**.28**	**.27**	**.27**	**.22** / **.15**	**.12** / .09
14. W2 Cat. Switch	**.10**	.07	**.13**	**.13**	**.11**	.08	**.29**	**.13**	**.12**	**.11**	**.17**	**.14**	**.15**	**.17** / **.11**

Note: Phenotypic correlations are displayed below the diagonal and Twin 1 – Twin 2 correlations are displayed above the diagonal. Correlations for MZ twins are displayed in the left and correlations for DZ twins are displayed on the right. Significant correlations are displayed in bold ($p < .05$). W1 = First wave of assessment (mean age 56 years); W2 = Second wave of assessment (mean age 62 years).

Table 3

Residual Genetic and Environmental Loadings and Correlations from the Longitudinal Model

Task	Wave 1 Residual ACEs			Wave 2 Residual ACEs			Cross-Wave Residual Correlations			
	A	C	E	A	C	E	r_A	r_C	r_E	r_{pheno}
Stroop	0.16	0.03	**0.69**	0.22	0.02	**0.65**	0.94	1.00	**0.29**	**0.44**
AX-CPT	0.10	0.07	**0.62**	0.03	0.21	**0.50**	0.53	1.00	**0.36**	**0.46**
Letter–Number	0.09	0.01	**0.40**	0.05	0.00	**0.53**	0.61	−1.00	**0.21**	**0.25**
Reading Span	**0.16**	**0.21**	**0.28**	**0.20**	0.00	**0.32**	**0.86**	1.00	**0.31**	**0.46**
Digit Span	0.04	0.00	**0.25**	0.05	0.01	**0.30**	1.00	−1.00	**0.20**	**0.32**
Trail-Making Test	0.13	0.00	**0.60**	0.02	0.00	**0.71**	1.00	−0.39	0.01	0.08
Category Switching	0.08	0.00	**0.86**	0.12	0.00	**0.81**	1.00	1.00	0.16	**0.25**

Note: Standardized variance components for the residual genetic (A), shared environmental (C), and nonshared environmental (E) influences on the seven EF tasks in the full longitudinal model (displayed in Figure 2). These residual variance components account for the remaining variation in each task not captured by the latent variables. The correlations between these residual components are also displayed, including the total estimated residual phenotypic correlation (final column). See supplement Table S2 for unstandardized estimates. Significant factor loadings and correlations are displayed in bold ($p < .05$).

Table 4

Model Comparisons for Tests of Factorial Invariance in the Longitudinal Model

Model	−2LL	df	BIC	RMSEA	vs. Model 1			vs. Model 2		
					diff (LL)	df	p	diff (LL)	df	p
1. Saturated ACE Cholesky	40697.87	16315	−67503	0.027						
2. Longitudinal Model – Configural invariance (Figure 2)	40835.83	16533	−68811	0.016	137.96	218	0.999			
3. Weak factorial invariance	40857.34	16541	−68843	0.017	159.47	226	0.999	21.51	8	0.006
4. Strong invariance (Figure 3)	40921.72	16546	−68811	0.022	223.85	231	0.655	85.89	13	< .001
5. Strict invariance	41022.55	16553	−68744	0.028	324.68	238	0.133	186.72	20	< .001

Note: Models 3–5 are subsets of the Longitudinal Model (Model 2, displayed in Figure 2) after equating unstandardized factor loadings (Model 3: weak factorial invariance), equating factor loadings and intercepts (Model 4: strong invariance), or equating factor loadings, intercepts, and residual variances (Model 5: strict invariance). Columns 6–8 show model comparisons with the saturated Cholesky model (Model 1) and columns 9–11 compare models to the longitudinal model (Model 2). −2LL = negative two times the log likelihood, BIC = Bayesian Information Criterion, RMSEA = Root Mean Square Error of Approximation.