Individual Differences in Executive Functions Are Almost Entirely Genetic in Origin

Naomi P. Friedman, Akira Miyake, Susan E. Young, John C. DeFries, Robin P. Corley, and John K. Hewitt
University of Colorado at Boulder

Recent psychological and neuropsychological research suggests that executive functions—the cognitive control processes that regulate thought and action—are multifaceted and that different types of executive functions are correlated but separable. The present multivariate twin study of 3 executive functions (inhibiting dominant responses, updating working memory representations, and shifting between task sets), measured as latent variables, examined why people vary in these executive control abilities and why these abilities are correlated but separable from a behavioral genetic perspective. Results indicated that executive functions are correlated because they are influenced by a highly heritable (99%) common factor that goes beyond general intelligence or perceptual speed, and they are separable because of additional genetic influences unique to particular executive functions. This combination of general and specific genetic influences places executive functions among the most heritable psychological traits. These results highlight the potential of genetic approaches for uncovering the biological underpinnings of executive functions and suggest a need for examining multiple types of executive functions to distinguish different levels of genetic influences.

Keywords: heritability, executive control, inhibition, task switching, frontal lobe tasks

Individuals vary widely in their abilities to control their own thoughts and actions. Some people seem ruled by impulses, whereas others manage successfully to regulate their behaviors. From the perspective of cognitive psychology, such variation reflects individual differences in executive functions, a collection of correlated but separable control processes that regulate lower level cognitive processes to shape complex performance. Although there is no clear consensus yet on how best to define or conceptualize executive functions, they are considered key mechanisms in many models of normal and abnormal cognition, such as cognitive development (e.g., Lyon & Krasnegor, 1996; Zelazo, Carter, Reznick, & Frye, 1997), age-related decline in cognitive abilities (e.g., Hasher, Zacks, & May, 1999; Lowe & Rabbitt, 1997), and disorders such as attention-deficit/hyperactivity disorder (ADHD; Barkley, 1997; Nigg, 2006), autism (Russell, 1997), schizophrenia (Frith, 1992), and substance use problems (Garavan & Stout, 2005).

Despite the centrality of executive functions in current psychological research, little is known about the sources of normal individual differences in executive functions. Why do people vary in their executive control abilities? We approached this question from a behavioral genetic perspective, which elucidates the etiology of individual differences by providing estimates of the extent to which they are due to genetic and environmental influences.

The Unity and Diversity of Executive Functions

Executive control has long been considered a unitary, general purpose ability that can be measured with a single complex "frontal lobe" task such as the Wisconsin Card Sorting Test (Heaton, Chelune, Talley, Kay, & Curtis, 1993). Recent behavioral and neuropsychological evidence indicates, however, that executive control may be more accurately characterized as a collection of related but separable abilities (Baddeley, 1996; Collette et al., 2005; Friedman et al., 2006), a pattern referred to as the "unity and diversity" of executive functions (Duncan, Johnson, Swales, & Freer, 1997; Miyake et al., 2000; Teuber, 1972). Researchers often disagree on what the underlying components of executive functions might be, but arguably the three most frequently studied executive functions are response inhibition (the ability to inhibit dominant, automatic, or prepotent responses), updating working memory representations (the ability to monitor incoming information for relevance to the task at hand and then appropriately update by replacing old, no longer relevant information with newer, more relevant information), and set shifting (the ability to flexibly switch back and forth between tasks or mental sets). Other executive functions have been examined, such as dual tasking (e.g., Logie, Cocchini, Della Sala, & Baddeley, 2004; Salthouse, Atkinson, & Berish, 2003) and resisting proactive interference (Friedman & Miyake, 2004), but the three executive functions mentioned above have dominated recent executive function research.

Understanding the structure of executive functions is complicated by the so-called task impurity problem (Burgess, 1997; Phillips, 1997): Because executive functions by definition operate on other cognitive processes, a large portion of the variance in any one executive function task is not necessarily measuring the putative executive process. Consider, for example, the Wisconsin Card Sorting Test, perhaps the most frequently used executive test in neuropsychological and molecular genetic studies. This task requires sorting cards along a particular dimension (e.g., the color, shape, or number of items on the cards), then switching to a different sorting dimension when the experimenter changes the correct sorting category (unbeknownst to the subject). In this case, performance reflects not only the ability to switch mental set (Miyake et al., 2000) but also perceptual, motor, and other cognitive abilities needed to sort the cards and monitor verbal feedback from the experimenter. This task impurity problem complicates the interpretation of studies based on a single complex executive function task, because it is unclear to what extent both null and positive results reflect such nonexecutive variance.

One method for alleviating the task impurity problem is the use of latent variables as dependent measures. Conceptually, a latent variable is an underlying ability that influences performance on a set of observed tasks, which are impure measures of this construct. It is estimated through a statistical extraction of the variance shared by multiple exemplar tasks selected to have different nonexecutive requirements but to tap the same underlying executive control ability. The resulting latent variable is a purer measure of this target ability and is virtually free from measurement error (Bollen, 1989).

Using this approach, Miyake et al. (2000) demonstrated that the three executive functions they examined—prepotent response inhibition (*Inhibiting*), updating working memory (*Updating*), and set shifting (*Shifting*)—were moderately correlated (i.e., showed unity) but were separable (i.e., showed diversity) at the level of latent variables. This general pattern of unity and diversity has since been replicated in other samples including young adults (Friedman et al., 2006), older adults (Fisk & Sharp, 2004; Hedden & Yoon, 2006), children (Huizinga, Dolan, & van der Molen, 2006; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; van der Sluis, de Jong, & van der Leij, 2007), and clinical populations selected for problems such as ADHD (Willcutt et al., 2001). Recent neuroimaging studies also indicate unity and diversity of executive functions in terms of brain localization (Collette et al., 2005; Sylvester et al., 2003). For example, Collette et al. used positron emission tomography to examine the brain areas that are common to as well as unique to inhibiting, updating, and shifting. These researchers used multiple tasks for each of the three executive functions, hence implementing the equivalent of latent variable analysis in a brain mapping context. They found common frontal and parietal areas activated by all three executive functions as well as frontal and/or posterior areas unique to updating and shifting.

The finding that executive functions show both unity and diversity has had important methodological and theoretical implications for many domains of psychological research. In particular, treating executive control as a multicomponent construct has enabled increased specificity about the nature of executive involvement in various cognitive, neuropsychological, and clinical constructs. In the cognitive domain, for example, Miyake et al. (2000) found that three complex neuropsychological and cognitive measures com-

monly used to assess general executive control ability were in fact differentially related to the three executive functions among young adults: the Wisconsin Card Sorting Test (most closely related to Shifting), random number generation (related to Inhibiting and Updating), and the Tower of Hanoi (related to Inhibiting). More recently, in an earlier subset of the sample used in the current study, Friedman et al. (2006) found that these three executive functions were differentially related to intelligence, with Updating, but not Inhibiting or Shifting, closely related to both fluid and crystallized intelligence.

In the clinical domain, considering the multiple components of executive functions has led to a better specification of the nature of executive deficits associated with various clinical disorders. For example, Willcutt et al. (2001) found that children with ADHD, reading disorders, and both disorders combined showed different profiles of executive dysfunction, with the ADHD children showing inhibiting deficits but the reading disorder children showing verbal working memory deficits (closely related to updating) and the ADHD with reading disorder group showing deficits on inhibiting, working memory, and shifting. Geurts, Verté, Oosterlaan, Roeyers, and Sergeant (2004) also found that children with ADHD were impaired on inhibiting and verbal fluency, whereas children with high functioning autism had more general executive function impairments. Moreover, data from the sample examined in the current study (unselected for cognitive or behavioral problems) suggest that Inhibiting is more closely related than Updating or Shifting to depressive symptoms (Sabella, Miyake, Friedman, Young, & Hewitt, 2007), externalizing behavior such as drug use and conduct disorder (Young et al., 2007), and classroom attention problems throughout childhood and adolescence (Friedman et al., 2007).

Thus, a wealth of studies illustrate the predictive validity of the inhibiting, updating, and shifting constructs for various cognitive abilities and real-world problems and the power afforded by measuring these constructs with latent variables. They also highlight the importance of considering executive control as a multidimensional construct when examining how it may be related to the abilities or problems of interest.

Specifying the Etiology of Cognitive Abilities With the Twin Design

Although these recent advances have increased our knowledge of the structure of executive functions, the etiology of this structure remains unspecified. Why do people vary in executive control abilities, and what makes these three executive functions related but separable? Twin designs using data from monozygotic (MZ; identical) and dizygotic (DZ; fraternal) twins can be used to estimate the extent to which interindividual variation in executive functions is influenced by additive genetic (A; heritability), common or shared environmental (C), and nonshared environmental (E) influences. Additive genetic influences for a complex trait are assumed to include the effects of a large number of specific genes that together operate in an additive manner. Shared environmental influences are those that contribute to similarity of twins (e.g., family environment, shared peer groups, mother's nutrition and hormone levels during gestation). In contrast, nonshared environmental influences are those that make twins' performances uncorrelated (e.g., different scholastic experiences, or even the same

scholastic environment if the twins respond differently to it for nongenetic reasons).

The basic logic for twin analyses is based on the fact that MZ twins share all of their genes whereas DZ twins share on average half of their genes by descent, and that both types of twins are reared together (i.e., have shared familial environment). Hence, MZ correlations within twin pairs that are higher than DZ correlations for a behavioral measure suggest a genetic influence on that measure. Specifically, given an additive genetic effect, one would expect an MZ correlation roughly twice the DZ correlation. When the DZ correlation is more than half the MZ correlation (i.e., when DZs show a higher correlation than would be expected given their genetic relatedness alone), shared environment is implicated. When MZ correlations are less than 1.0 (i.e., MZs show lower correlations than would be expected if the trait were due entirely to genetic and shared environmental influences), nonshared environment is implicated. Note that at the level of individual tasks, nonshared environment can also include measurement error, because such error will tend to make twins' performances uncorrelated.

To decompose the variance in a dependent measure due to these genetic and environmental (shared and nonshared) influences, behavior geneticists typically use an *ACE* structural equation model (Neale & Cardon, 1992). Figure 1 illustrates the general form of the ACE model for a single observed measure (known as a phenotype), measured in MZ and DZ twins. The rectangles denote the phenotype (such as IQ) measured in both twins in a pair. The three circles (A, C, and E) represent the latent (unobserved) variance components that are estimated from the MZ and DZ data. The model is set up as a two-group analysis (one for MZ twins, the other for DZ twins), with the estimated A, C, and E variances constrained to be equal across groups. Several constraints are

Figure 1. General ACE structural equation twin model of a single behavioral measure. Individual differences in performance on a measure (depicted with rectangles to denote an observed variable) are modeled as due to three types of influences: additive genetic (A), shared environmental (C), and nonshared environmental (E). The Twin 1 with Twin 2 A correlation is set to 1.0 for monozygotic (MZ) twins because they share all of their genes and 0.5 for dizygotic (DZ) twins because they share on average half of their genes by descent. The C correlation is set to 1.0 for both types of twins because both are raised together. The E correlations are set to zero because nonshared environment is uncorrelated by definition.

implemented to estimate these A, C, and E variances: The correlation between the genetic effects (A) in the MZ twins is set to 1.0, because they share all of their genes. This correlation in the DZ twins is set to 0.5, as they share on average half of their genes by descent. Because both types of twins are reared together, the correlation between shared environmental influences (C) is set to 1.0 in both groups. All correlations with nonshared environmental influences (E) are set to zero, because they are unique by definition. Fitting this model to the covariance matrices of task performance for the MZ and DZ twins (i.e., the covariance matrices relating Twin 1's performance to Twin 2's performance in each zygosity group) provides estimates of these three effects.

In the domain of cognitive abilities, such models have been used to decompose the variances in general cognitive ability (*g*) and specific cognitive abilities (e.g., verbal and spatial abilities, memory, and speed) into their respective genetic and environmental components. By and large, twin studies (along with family and adoption studies) indicate that *g*, whether measured as the variance common to numerous tasks or with an assessment of general IQ, is moderately heritable: generally about 50%, with the remaining environmental variance split fairly evenly between shared and nonshared (e.g., Chipuer, Rovine, & Plomin, 1990; Neisser et al., 1996). These estimates, however, change across development, with the role of shared environment in IQ decreasing to almost zero and the heritability increasing to 70% or greater by late adolescence (McGue, Bouchard, Iacono, & Lykken, 1993; Neisser et al., 1996). Studies of specific cognitive abilities suggest that they are also moderately heritable (about 30% to 60%) by late adolescence and into adulthood (Alarcón, Plomin, Fulker, Corley, & DeFries, 1999; Pedersen, Plomin, Nesselroade, & McClearn, 1992), with the remaining variance largely due to nonshared environment.

The basic ACE model can be extended to multivariate analyses to address questions beyond the etiology of individual differences in a single ability. In particular, multivariate analyses can be used to examine the nature of the relationships among different constructs by specifying the extent to which they share genetic and environmental influences and which types of influences (genetic and/or environmental) differentiate them. For example, such multivariate models have been useful for specifying the structure of specific cognitive abilities and their relation to *g*. The picture that has emerged (Alarcón et al., 1999; Petrill, 1997) is that the moderate genetic effects observed for different specific cognitive abilities (e.g., verbal, spatial) are largely general (i.e., influence multiple cognitive abilities), with only small genetic influences unique to each type of specific cognitive ability. It is primarily environmental influences that differentiate one specific cognitive ability from another.

The Current Study

In the current study, we used such ACE models to answer two questions regarding the etiology of executive functions: (a) How much of the individual differences in these executive functions are due to genetic and environmental influences? and (b) Do genetic influences on these three executive functions operate at only the general level (i.e., common to all three executive functions) or at both general and specific levels?
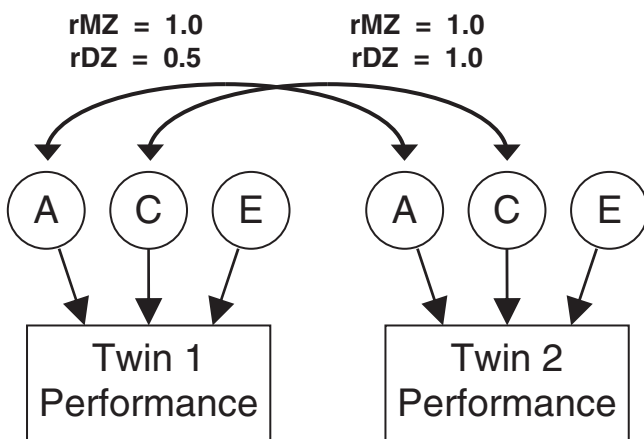
A few previous twin studies have examined the first question at the level of individual tasks in adolescents or adults. These studies generally suggest moderate genetic influence (around 40% to 60% on average), virtually no significant shared environmental influence, and moderate nonshared environmental influence on individual tasks thought to tap inhibiting, updating, and shifting abilities, such as the antisaccade task, working memory span tasks, and the Wisconsin Card Sorting Test (e.g., Ando, Ono, & Wright, 2001; Anokhin, Heath, & Myers, 2004; Anokhin, Heath, & Ralano, 2003; Fan, Wu, Fossella, & Posner, 2001; Luciano et al., 2001; Malone & Iacono, 2002; Posthuma, Mulder, Boomsma, & de Geus, 2002; Wright et al., 2001).

Although these studies provide preliminary evidence that executive functions may be somewhat genetically influenced, the task impurity problem complicates the interpretation of heritability estimates based on single complex executive function tasks, because it is unclear to what extent those results reflect nonexecutive variance. As discussed earlier, one fruitful method for circumventing this problem in behavioral and neuropsychological studies is to adopt a latent variable approach, but this method has not been applied to multivariate genetic studies targeting specific executive function constructs. The use of latent variables in a genetic context is highly informative, because it enables separation of the genetic and environmental influences on each task into those influencing the target executive function and those influencing the nonexecutive components of the tasks. For example, if the heritability estimates for the individual tasks for a particular executive function reflect only the variance associated with nonexecutive requirements of those tasks, then the variance shared among these tasks should not be heritable at all. Perhaps a more likely possibility is that the genetic influences on multiple executive function tasks overlap somewhat, which would result in some heritability of the latent variable, though there could also still be some genetic influences on the nonexecutive components of each task. As these examples illustrate, by conducting behavioral genetic analyses at the level of latent variables, with the current study we seek to provide more precise estimates of the genetic and environmental influences on individual differences in executive functions than have to date been provided.

An examination of multiple executive functions at the level of latent variables also enables us to address the second research goal above, namely to specify the level (or levels) at which genetic influences on executive functions operate and thereby shed new light on the etiology of the unity and diversity of executive functions. For example, one plausible scenario is that executive functions are related because they share common genetic and environmental influences but are separable purely because of environmental influences unique to individual executive functions. In this case, there would be only one level of genetic influences (the general level that is common to all three executive functions). Those influences could overlap considerably with those that affect g, which is known to have moderate to large genetic (50% to 75%) as well as some environmental (mostly nonshared) influences by late adolescence (McGue et al., 1993; Neisser et al., 1996).

Another possible configuration that could explain the unity and diversity of executive functions is that the unity is due to shared genetic and environmental influences, as in the previous scenario, but the diversity is also due to genetic and environmental influences unique to the individual executive functions (rather than just environmental influences). In this case, there are two levels of genetic influences: one that operates at the general level, producing unity, and another that operates on specific executive functions, producing diversity. This configuration would be similar to that found for specific cognitive abilities, which seem to be related through genetic and environmental influences on g and separable because of large environmental and, in some cases, small genetic influences unique to particular cognitive abilities (Alarcón et al., 1999; Petrill, 1997).

Of course, many other combinations of genetic and environmental influences could logically explain the unity and diversity of executive functions. Hence, the "problem space" for understanding individual differences in executive functions is large. Reducing this space is important, because it will likely constrain what lines of research will be most productive for further understanding the nature of executive functions.

For example, a good deal of research on executive control has begun to incorporate genetic information with the aim of understanding the neurobiology of executive functions and, by extension, disorders often associated with impairments in executive control. This line of research has focused on finding specific genetic variants associated with executive control abilities and/or related disorders such as ADHD (e.g., Faraone et al., 2005) and schizophrenia (e.g., Egan et al., 2001). Although such genetic association studies have great potential for elucidating the genetic and biological bases of executive functions, the lack of a clear picture of the underlying genetic and environmental structure of executive functions limits their impact. In particular, most molecular genetic studies have focused on gross executive control ability as measured by individual neuropsychological tests such as the Wisconsin Card Sorting Test. However, the emerging evidence that there are separable executive functions and that those executive functions differentially relate to disorders of interest begs the question of whether such research should incorporate multiple executive functions and multiple tasks for each function. The answer to this question depends on whether there are genetic influences on executive functions at the level of latent variables and, if so, whether those influences operate at only the general level (i.e., common to all executive functions) or at both the general and specific levels.

In the current study, we took a latent variable approach to answer the two primary questions. To specify the etiology of executive functions, we estimated the ACE components of each executive function, measured as a latent variable. To specify the etiology of the unity and diversity of executive functions, we then used a multivariate ACE model of all three executive functions, estimating ACE components for both what is common to the three executive functions and what is unique to each executive function. Once we specified the genetic and environmental structure of these three executive functions, we tested two key alternative hypotheses regarding the unity of executive functions: that it primarily reflects the same genetic and/or environmental influences as speed or IQ. More specifically, in these secondary analyses, we examined to what extent executive function variance overlaps with variance in IQ and speed, respectively.

Method

Participants

Participants were 582 individuals from 293 same-sex twin pairs, recruited from the Colorado Longitudinal Twin Study. The sample consisted of 316 MZ (177 female, 139 male) twins and 266 DZ (137 female, 129 male) twins (4 co-twins did not participate in either the year 16 or year 17 testing). The current Longitudinal Twin Study sample consists of approximately 420 families that met certain criteria (twins with normal birth weights, gestation periods, and a residence located within 2 hr of Boulder, Colorado) who agreed to participate after being located through birth records provided by the Division of Vital Statistics of the Colorado Department of Health from 1986 through 1990 (for more detailed information about this longitudinal sample, see Rhea, Gross, Haberstick, & Corley, 2006). The analyses for this study were conducted on a subset of families in this sample that had completed the executive function battery. These twins are representative of the general population in terms of IQ, as indicated by a normal IQ distribution ($M = 102$, $SD = 11$). Twins in each pair were randomly assigned to Twin 1 and Twin 2 for the analyses. Participants received $50 compensation for the approximately 3-hr session. The protocol and informed consent procedures were approved by the University of Colorado Human Research Committee (Protocol 0600.01).

Zygosity was determined through repeated tester ratings combined with DNA genotyping. First, testers' ratings of each twin pair on a nine-item physical characteristic assessment (Nichols & Bilbro, 1966) were used to judge zygosity (the mean number of judgments for the 293 pairs was 18.9 for this longitudinally followed sample; the range was 2 to 30). Second, using DNA from cheek swabs, twins were genotyped at a minimum of 11 informative short tandem repeat polymorphisms (STRPs) using standard polymerase chain reaction methods and ABI 377 genotyping technology. Concordance for all STRPs between twins indicated MZ status. DNA was available for 291 twin pairs; of these, genotyping resulted in unambiguous zygosity calls for 280 of the twin pairs. For these 280 pairs, any pairs with discrepancies between the ratings and DNA calls were evaluated by senior staff, and if necessary, DNA was resampled and regenotyped.

Materials, Design, and Procedure

The nine executive function tasks were based primarily on those we have used successfully before (Miyake et al., 2000). All nine executive function tasks were computerized (Macintosh iBook computers) in PsyScope 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993). A button box with millisecond accuracy was used to measure reaction times (RTs), and a headset was attached to the button box to record RTs for verbal responses. Stimuli within tasks were appropriately counterbalanced and randomized, and the order of stimuli within each task was the same for all participants. All tasks included additional practice trials to ensure that the participants fully understood the instructions and that they had firmly mastered the button–response mappings.

Inhibiting Tasks

Antisaccade. During each trial of the antisaccade task (adapted from Roberts, Hager, & Heron, 1994), a fixation point appeared in the center of the computer screen for a variable amount of time (one of nine times between 1,500 and 3,500 ms in 250-ms intervals). A visual cue (a 1/8-in. [0.32-cm] black square) then appeared on one side of the screen for 150 ms, followed by the target stimulus (a 5/16-in. [0.79-cm] arrow inside of an open 7/16-in. [1.11-cm] square) on the opposite side for 175 ms. The target was then masked with gray cross-hatching, and the mask remained on the screen until the participant indicated the direction of the arrow (left, up, or right) with a button press response. The inner edges of the cues and targets were 3 in. (7.62 cm) and 3.625 in. (9.21 cm), respectively, away from the fixation point (on opposite sides). The participants were seated 18 in. (45.72 cm) from the computer monitor. The participants practiced on 22 trials and then received 90 target trials. The dependent measure was the proportion of correct responses.

Stop signal. The stop-signal task (Logan, 1994) consisted of five blocks of trials. On each trial in the first block of 48 trials, used to build up a prepotent categorization response and calculate each participant's average RT, participants saw one of 24 words (e.g., *duck*, *gun*) and categorized it as either an animal or a nonanimal as quickly as possible without making mistakes. Then, in the four subsequent blocks of 96 trials each, participants tried *not* to respond (i.e., to inhibit the categorization response) when they heard a computer-emitted signal (a tone approximately 100 ms long) on a randomly selected 25% of the trials, but otherwise kept performing the same categorization task. In all trials (including 34 practice trials, 24 before the first no-signal block and 10 at the beginning of the first signal block), the participants viewed a fixation point for 500 ms and were then allowed up to 1,500 ms to categorize the target word. Each participant experienced signals that occurred 50 ms before his or her average RT (long stop-signal delay), 225 ms before his or her average RT (medium stop-signal delay), or 50 ms after the onset of the trial (short stop-signal delay). Each of these delays occurred equally often in each block. As recommended by Logan (1994), the instructions emphasized that the participants should not slow down to wait for possible signals. The dependent measure was the stop-signal RT, the estimated time at which the stopping process finishes. We used the most common estimation method, which assumes that the stop-signal RT is a constant (Logan, 1994). Specifically, the stop-signal RT for each delay was calculated as follows: The RTs for the no-signal go trials were rank ordered, and the stop-signal delay was subtracted from the *n*th RT, where *n* is the number of all of the no-signal RTs multiplied by the probability of responding at that delay. Then the stop-signal RTs for all delays were averaged.

Stroop. On each trial of the Stroop task (Stroop, 1935), adapted for computer administration, participants saw a white fixation point on a black screen for 500 ms, followed by the stimulus, which remained on the screen until the participant responded, after which the screen remained black for 1,000 ms. Participants verbally named the color of each stimulus as quickly and as accurately as possible, with RTs measured by voice key. There were three types of trials: (a) 60 trials with a string of asterisks (of variable lengths matching the lengths of the color words) printed in one of six colors (red, green, blue, orange, yellow, or purple); (b) 60 trials with a color word printed in a different color (e.g., *BLUE* printed in red); and (c) 60 filler trials (not used in the current study) with a neutral word printed in one of the six colors. The different trial types were not blocked. The

order of the trials was randomized with the constraints that no word or color on one trial was related to the word or color on the immediately preceding trial and that no condition appeared more than three trials in a row. The trials were broken down into four subblocks. The participants also received voice-key calibration and 18 practice trials. The dependent measure was the RT difference between the trials in which the word and the color were incongruent and the trials that consisted of asterisks.

Updating Tasks

Keep track. In each trial of the keep track task (adapted from Yntema, 1963), participants were first shown several target categories at the bottom of the computer screen. Fifteen words, including two or three exemplars from each of six possible categories (animals, colors, countries, distances, metals, and relatives), were then presented serially and in random order in the center of the screen for 1,500 ms each, with the target categories remaining at the bottom of the screen. The task was to remember the last word presented in each of the target categories and then report these words at the end of the trial. For example, if the target categories were metals, relatives, and countries, then at the end of the trial, participants recalled the last metal, the last relative, and the last country presented in the list. During the instruction period, participants saw all six categories and the exemplars in each to ensure that they knew to which category each word belonged. They practiced on 3 trials (progressing in difficulty from two to four target categories), then performed 12 trials (four of each difficulty, presented in random order), recalling a total of 36 words. The proportion of words recalled correctly was the dependent measure.

Letter memory. In the letter memory task (adapted from Morris & Jones, 1990), several letters from a list were presented in the center of the screen serially for 2,500 ms per letter. The task was to recall the last three letters. To ensure that the task required continuous updating, the instructions required the participants to rehearse out loud the last three letters by mentally adding the most recent letter and dropping the fourth letter back, and then saying the new string of three letters out loud. For example, if the letters presented were *T, H, G, B, S, K, R*, the participants should have said, "T . . . T-H . . . T-H-G . . . H-G-B . . . G-B-S . . . B-S-K . . . S-K-R" and then recalled "S-K-R" at the end of the trial. The number of letters presented (five, seven, or nine) was varied randomly across trials, with the constraint that each list length was used once in every three trials. Each list began with a 1-s fixation point, which was followed by the letters and then a string of three pink question marks that remained on the screen until the participant finished recalling the target letters. Participants were instructed to recall the letters in order and to say "blank" if they did not remember a particular letter. However, answers were scored as correct even if the letters were not recalled in the correct order. After practicing on 3 trials (one of each length), the participants completed 12 trials (four of each length). Because of a programming error, however, the last two lists (five and seven letters long) were unusable. Hence, only the first 10 trials were used (30 letters recalled). The dependent measure was the proportion of letters correctly recalled across all lists.

Spatial 2-back. In each block of the spatial 2-back task, there were 10 open 5/8-in. (1.59-cm) squares scattered across the screen.

After a beep, one box at a time became solid black for 500 ms, giving the appearance that it flashed. There were 1,500 ms between each flash (24 flashes per block). For each flash, participants pressed a button indicating whether that box was the same one that had flashed two trials earlier (there were six "yes" flashes in each block). There were no instances in which the current flash was the same as the one that was one or three trials back. Participants completed one practice block and then four actual blocks, with breaks between each block (participants pressed a button to begin each new block). The dependent measure was the proportion of correct responses (yes and no) across all four blocks. Omissions were counted as errors.

Shifting Tasks

Three tasks were used to assess Shifting ability. In each of the three tasks, there were four blocks of 48 trials, each of which contained 24 no-switch and 24 switch trials. Each trial was preceded by a cue indicating which subtask should be performed on that trial, and the cue remained on the screen throughout the trial. In the first and third blocks, the cue was presented 150 ms before the onset of the stimulus, and both the cue and the stimulus remained on the screen until the participant responded, at which point the next cue appeared after a 350-ms response-to-cue interval. In the second and fourth blocks (not analyzed in the current study), everything was the same except that the cue appeared 1,500 ms before the onset of the stimulus. Throughout each task, participants were asked to use whatever time they had between the cue and the stimulus to prepare for the forthcoming subtask. They were also asked to respond as quickly as possible without making mistakes. To firmly master the cue–subtask associations and the key mappings, participants completed two practice blocks of 24 trials each before the task began. In addition, there were 6 warm-up trials at the beginning of each block that were not analyzed. For all tasks, the order of the trials was randomized with the constraint that no more than 4 switch trials could occur in a row. Further, there were no item-specific negative priming trials in which the stimulus on a switch trial was the same as that on the previous trial (except for the color–shape task, for which such trials were unavoidable due to the small number of potential stimuli). The dependent measure in each task was the regular switch cost, calculated as the difference between the average RTs of the trials that required a switch and the average RTs of the trials in which no switch was necessary for the trials with the short (150-ms) cue-to-stimulus interval.

Number–letter. In each trial of this task (adapted from Rogers & Monsell, 1995), a number–letter or letter–number pair (e.g., *7G*) was presented in one of two squares above or below a line dividing the computer screen in half. The participants were instructed to indicate whether the number was odd or even (*2, 4, 6,* and *8* for even; *3, 5, 7,* and *9* for odd) when the pair was in the top square and to indicate whether the letter was a consonant or a vowel (*G, K, M,* and *R* for consonant; *A, E, I,* and *U* for vowel) when the pair was in the bottom square. The cue in this task was the onset of the square. The squares were 3/4 in. (1.91 cm) square, and their edges appeared 3/16 in. (0.48 cm) above or below the median line.

Color–shape. In each trial of this task (Miyake, Emerson, Padilla, & Ahn, 2004), a cue letter (*C* or *S*) appeared above a colored rectangle with a shape in it (outline of a circle or triangle).

The participants were instructed to indicate whether the color was red or green when the cue was *C* and whether the shape was a circle or triangle when the cue was *S*. The colored rectangles were approximately 1.7 in. (4.32 cm) wide and 1.4 in. (3.56 cm) high, the circles were approximately 1.1 in. (2.79 cm) in diameter, and the triangles were 1.25 in. (3.18 cm) on each side. The color–shape figure appeared in the center of the screen, and the cue letter was centered 3/8 in. (0.95 cm) above its top edge.

Category switch. In each trial of this task (adapted from Mayr & Kliegl, 2000), participants saw a word that could be categorized in terms of (a) whether it described a living or nonliving thing or (b) whether it described a thing that is smaller or larger than a soccer ball. The 16 words were drawn from those used by Mayr and Kliegl: *table*, *bicycle*, *coat*, *cloud*, *pebble*, *knob*, *marble*, *snowflake*, *shark*, *lion*, *oak*, *alligator*, *mushroom*, *sparrow*, *goldfish*, and *lizard*. A symbol appearing above the word cued which categorization to use (a heart indicated living vs. nonliving and an arrow cross indicated large vs. small). The words were presented in the center of the screen, and the 9/16 in. (1.43 cm) high by 11/16 in. (1.75 cm) wide symbols appeared 3/8 in. (0.95 cm) above them.

General Cognitive Ability

The Wechsler Adult Intelligence Scale—Third Edition (WAIS–III; Wechsler, 1997), a test of general intelligence, consists of 11 subtests (Vocabulary, Similarities, Arithmetic, Digit Span, Information, Comprehension, Picture Completion, Digit Symbol, Block Design, Picture Arrangement, and Object Assembly). Scaled scores from all subtests were used to compute a full-scale IQ score.

Perceptual Speed

Hidden patterns. In this test (Ekstrom, French, Harman, & Derman, 1976), participants viewed a simple criterion or model pattern. Then, for a series of slightly more complex test patterns, they determined whether the criterion pattern was embedded in each test pattern. There were 200 items in each of the two subsections (the criterion pattern remained the same throughout the entire test), and the time limit was 2 min for each subsection. The dependent measure was the number correct minus the number incorrect (correction for guessing).

Perceptual speed. Each of the two parts of the Colorado Perceptual Speed Test (DeFries, Plomin, Vandenberg, & Kuse, 1981) contained 30 items consisting of a stimulus of letters or numbers and a set of four alternative responses. Participants were allowed 1 min in each part to find the responses that were identical to the stimuli. The dependent measure was the number correct minus 1/3 times the number incorrect (correction for guessing).

Identical pictures. This test (Ekstrom et al., 1976) required participants to view a target figure and judge which one of five alternative figures was identical to the target figure as quickly and as accurately as possible. There were 48 items in each of the two subsections, and the time limit was 1.5 min for each subsection. The dependent measure was the number of items correct minus 1/4 times the number of items incorrect (correction for guessing).

General Procedure

Participants completed the measures of intelligence and perceptual speed in prior testing sessions at approximately age 16 (*M* =

16.6 years, *SD* = 0.8, range = 15.8 to 20.0). They completed the executive function tasks at approximately age 17 (*M* = 17.3 years, *SD* = 0.6, range = 16.1 to 20.1). The tasks administered at age 17 were grouped into three blocks lasting approximately 40–50 min each, in which no two tasks intended to tap the same construct were presented sequentially. The order of task administration was fixed as follows for all participants to minimize any measurement error due to participant by order interaction: antisaccade, letter memory, color–shape, 5-min break, number–letter, Stroop, keep track, lunch break, spatial 2-back, category switch, and stop signal. Other measures included as parts of other studies (not analyzed in the current study) were also administered, the majority of which were completed after these three blocks of tasks.

Statistical Procedures

Data Trimming and Outlier Analyses

For the RT measures (except stop signal, which did not depend on a mean RT), RTs from errors (voice key or others) and RTs below 200 ms were eliminated (following our previous procedures; Miyake et al., 2000). For the three shifting tasks, RTs for trials immediately following errors were also excluded, because the correct set might not have been achieved on the prior error trials. Average accuracy was greater than 92% in all RT tasks. To obtain the best measure of central tendency for each condition in the tasks using RT difference scores, we applied a within-subject trimming procedure that is robust to nonnormality (Wilcox & Keselman, 2003): For each participant, observations that deviated from the median by more than 3.32 times the median absolute deviation in each condition were excluded.

All accuracy data were arcsine transformed to improve normality. To reduce the influence of extreme scores and improve normality, we replaced observations farther than three standard deviations from the group mean with values three standard deviations from the mean for each variable used in the models except WAIS–III full-scale IQ (WAIS–IQ). This procedure affected no more than 2.1% of the observations for any measure. After these transformations and trimming, the variables showed acceptable skewness and kurtosis (see Table 1). In all analyses, the directionality of the RT measures was reversed so that for all measures, higher scores indicated better performance.

Model Estimation

We used Mplus 4.0 (Muthén & Muthén, 2006) and Mx (Neale, Boker, Xie, & Maes, 2003) to estimate the latent variable models with maximum-likelihood estimation of the raw data, including participants with missing data. This approach provides the optimum use of all of the available twin data (Neale et al., 2003), providing unbiased estimates of parameters for the population where most other approaches (e.g., listwise deletion of individuals with missing data) would not (Little & Rubin, 1987). Eighty-seven participants were missing data for one or more executive function tasks because of colorblindness, equipment malfunction, failure to understand or follow task instructions, chance-level accuracy, or

Table 1
Descriptive Statistics

Task	N	M	SD	Min	Max	Skewness	Kurtosis	Reliability
Antisaccade[a]	562	1.04	0.20	0.47	1.57	−0.14	−0.32	.89[b]
Stop signal	540	284 ms	65	151	500	1.18	1.67	.76[b]
Stroop	548	212 ms	89	0	483	0.51	0.06	.91[b]
Keep track[a]	559	0.93	0.18	0.37	1.50	0.25	0.59	.66[c]
Letter memory[a]	568	1.09	0.24	0.38	1.57	0.29	−0.01	.61[c]
Spatial 2-back[a]	564	1.17	0.18	0.63	1.57	−0.96	1.57	.91[c]
Number–letter	562	336 ms	190	−14	953	1.12	1.28	.86[b]
Color–shape	551	333 ms	192	−196	930	0.83	0.97	.86[b]
Category switch	553	343 ms	193	−34	941	1.05	1.01	.85[b]
WAIS–IQ	582	102	11	70	142	0.16	0.37	.97[d]
Hidden patterns	572	90	23	20	161	−0.10	0.43	.91[b]
Perceptual speed	580	37	8	13	59	−0.03	0.09	.86[b]
Identical pictures	580	80	13	42	96	−0.58	0.49	.87[b]

Note. Min = minimum; Max = maximum; WAIS–IQ = Wechsler Adult Intelligence Scale full-scale IQ.
[a] Accuracy scores were arcsine transformed. [b] Internal reliability was calculated by adjusting split-half or Part 1–Part 2 correlations with the Spearman–Brown prophecy formula. [c] Internal reliability was calculated using Cronbach's alpha. [d] Internal reliability from Wechsler (1997).

other reasons [1] (see Table 1 for *n*s). Where appropriate (i.e., for the antisaccade, stop signal, number–letter, color–shape, and category switch tasks), the criterion for chance performance for each task was calculated as the binomial probability that the participant would have obtained that score by chance, with $p < .01$.

Because the conventional measure of fit, the chi-square (χ^2) statistic, is sensitive to sample size (Kline, 1998), we used $\chi^2/df < 2$ (Byrne, 1989) as an indication of good fit. We supplemented it with two other types of fit indices: the root-mean-square error of approximation (RMSEA) statistic, an absolute fit index that quantifies how closely the covariances predicted by the model match the actual covariances, and the Tucker–Lewis index (TLI), a nonnormed (i.e., can exceed 1.0) incremental fit index that quantifies how well the model fits compared with a null model (in this case, a model with only means and variances but no covariances among the measures). We selected these particular fit indices because they both compensate for the effect of model complexity by taking into account the degrees of freedom of the models (Hu & Bentler, 1998; Marsh, Balla, & Hau, 1996), an important consideration given the complexity of the multivariate genetic models. Following the recommendation of Hu and Bentler, we used RMSEA < .06 and TLI > .95 as indications of good fit, though TLI > .90 is commonly used as an indication of adequate fit.

To correct for the nonindependence of the twin pairs in the phenotypic (nongenetic) analyses, we used Mplus's TYPE = COMPLEX option to obtain corrected standard errors and a scaled chi-square robust to nonindependence. We used chi-square difference tests for nested model comparisons, appropriately scaled (Satorra & Bentler, 2001) for nonindependence in the phenotypic models. All analyses used an alpha level of .05. For the genetic models, we designated parameters as significant if their bootstrapped 95% confidence intervals (estimated with Mx) did not include zero, and in some key cases, we also checked significance with chi-square difference tests.

For the genetic analyses, we used a general form of the ACE model to estimate genetic and environmental effects (Neale & Cardon, 1992). Means for each task were constrained to be equal in all groups (i.e., there was one estimated mean for each task, which was constrained to be the same for Twin 1 and Twin 2 and across zygosity groups), because there was no reason to expect any differences.

Results and Discussion

The primary goals of the current study were to (a) specify the extent to which individual differences in the three target executive functions (Inhibiting, Updating, and Shifting) are due to genetic and environmental influences and (b) specify how these genetic and environmental influences combine to form the unity and diversity of executive functions. Toward these goals, we present two main genetic analyses: ACE models of the individual executive functions separately and a multivariate ACE model of the three functions together. We then present secondary genetic analyses that address whether the genetic variance common to the three executive functions goes beyond that for perceptual speed and *g*. Before doing so, however, we briefly present the results of nongenetic (phenotypic) confirmatory factor analyses to verify that the three executive functions show the same factor structure as in our previous investigation (Miyake et al., 2000).

[1] Included in this count of 87 were 13 participants for whom all executive function data were unusable because there were problems in the testing environment ($n = 1$), the participants were unable to understand instructions ($n = 2$), or the participants' data were suspect because of repeated chance-level performance (on at least four individual tasks; $n = 9$) or clear fatigue ($n = 1$). There was also one co-twin who did not participate in the executive function (age 17) testing session, but had participated in the earlier (age 16) cognitive testing session. These participants' executive function data were excluded from the analyses. However, their IQ and Perceptual Speed data from the earlier session were included in the analyses. Hence, the total sample size for the executive function models was 568 (from 158 MZ families and 134 DZ families), and the total sample size for the additional models that included the perceptual speed or IQ data was 582 (from 159 MZ families and 134 DZ families).

Phenotypic Analyses

Confirmatory Factor Analysis of the Three Executive Functions: Do They Show Unity and Diversity?

Descriptive statistics for the nine executive function tasks, WAIS–IQ, and the perceptual speed tasks are presented in Table 1,[2] and maximum-likelihood estimates of the correlations among all tasks are presented in Appendix A.[3] As shown in Figure 2, the three executive functions (depicted as ellipses to denote that they are latent) correlated moderately with one another (Inhibiting with Updating, $r = .74$; Inhibiting with Shifting, $r = .73$; Updating with Shifting, $r = .40$). The fit of this full, three-factor model was good, $\chi^2(24) = 44.51$, $p = .007$, RMSEA = .039, TLI = .955, according to our criteria of $\chi^2/df < 2$, RMSEA < .06, and TLI > .95. Comparisons of this model to restricted models in which each correlation was set to zero or 1.0 indicated that these correlations were significantly larger than zero, all $\chi^2_{diff}(1) > 24.88$, $p < .001$, but significantly smaller than 1.0, all $\chi^2_{diff}(1) > 11.30$, $p < .001$. Hence, each pair of latent variables was significantly correlated, but no pair of was perfectly correlated. Moreover, as shown in Table 2, the fit of the three-factor model was significantly better than the fits of reduced two-factor models in which two of the three latent variables were collapsed, all $\chi^2_{diff}(2) > 43.03$, $p < .001$. The three-factor model also fit much better than a one-factor model with all three executive functions collapsed into a single factor, $\chi^2_{diff}(3) = 220.36$, $p < .001$. Hence, these results replicate our previous finding of unity and diversity of these three executive functions (Miyake et al., 2000).

Also shown in Figure 2, the nine tasks loaded significantly on their respective factors, all $\chi^2_{diff}(1) > 51.42$, $p < .001$, with the factor loadings ranging from .42 to .74. Subtracting the square of each factor loading from 1.0 gives the residual variance in each task that is unrelated to the target executive function (i.e., task impurity as well as measurement er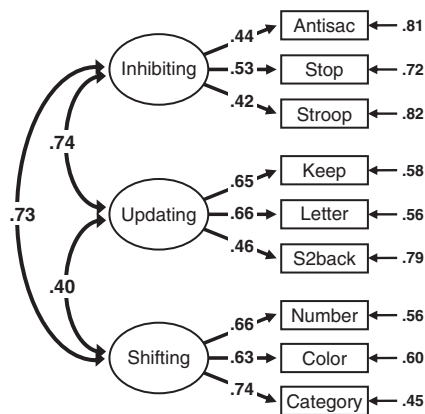ror). As shown on the far right side of Figure 2, these residual variances ranged from .45 to .82. Hence, each task had substantial variance that was unrelated to the executive functions of interest, underscoring the importance of using latent variables to separate out this task impurity.

Secondary Models: How Do the Executive Functions Relate to Speed and IQ?

In an additional confirmatory factor analysis, we added the Perceptual Speed latent variable (which had identical pictures, Colorado Perceptual Speed Test, and hidden patterns as indicators) and allowed it to correlate with the three executive functions. In the resulting model, $\chi^2(48) = 119.85$, $p < .001$, RMSEA = .051, TLI = .923, Perceptual Speed correlated .69, .55, and .47 with Inhibiting, Updating, and Shifting, respectively, again all significantly greater than zero, all $\chi^2_{diff}(1) > 45.55$, $p < .001$, but significantly smaller than 1.0, all $\chi^2_{diff}(1) > 16.05$, $p < .001$. These three correlations could not be constrained to be equal, $\chi^2_{diff}(2) = 7.51$, $p = .023$. These results are consistent with recent findings that executive functions are related to processing speed (e.g., Hedden & Yoon, 2006; Salthouse et al., 2003), though the magnitudes of the correlations were far smaller than unity in the current study.

In a third confirmatory factor analysis, we added WAIS–IQ as an observed variable to the model in Figure 2 and allowed it to correlate with the three executive functions. In this model, $\chi^2(30) = 63.01$, $p < .001$, RMSEA = .043, TLI = .944, WAIS–IQ correlated .53, .70, and .19 with Inhibiting, Updating, and Shifting, respectively. All of these correlations were significantly greater than zero, all $\chi^2_{diff}(1) > 10.90$, $p < .001$, but significantly smaller than 1.0, all $\chi^2_{diff}(1) > 37.90$, $p < .001$. These three correlations could not be constrained to be equal without worsening the model fit, $\chi^2_{diff}(2) = 26.19$, $p < .001$. This differential relationship between WAIS–IQ and the three executive functions—especially IQ's strong association with Updating—echoes the findings that we reported earlier based on a subset of the current sample (Friedman et al., 2006). Moreover, the finding that IQ is related relatively weakly to Shifting abilities is also consistent with some other recent studies (Süβ, Oberauer, Wittmann, Wilhelm, & Schulze, 2002; Yehene & Meiran, 2007).

Genetic Analyses

Task ACE Models: Etiology of Individual Differences in Individual Executive Tasks

As stated earlier, our primary interests in the genetic analyses were the ACE estimates for the latent executive function variables

Figure 2. Phenotypic confirmatory factor analysis model of the three executive functions. Numbers on arrows are standardized factor loadings, those next to the smaller arrows on the right are residual variances, and those on curved double-headed arrows are interfactor correlations. All parameters were significant ($p < .05$). Antisac = antisaccade; stop = stop signal; keep = keep track; letter = letter memory; S2back = spatial 2-back; number = number–letter; color = color–shape; category = category switch.

---

[2] Sex differences in executive function task performance were not significant, except for antisaccade (arcsined accuracy for males = 1.09 [SD = 0.18] and females = 1.00 [SD = 0.21], $p < .001$). Males and females did not significantly differ in their executive function factor structures, $\chi^2_{diff}(12) = 12.12$, $p = .436$; executive function latent variable means, all $\chi^2_{diff}(1) < 2.97$, $p > .085$; or univariate task ACE estimates, all $\chi^2_{diff}(2) < 4.07$, $p > .131$, except for stop signal, $\chi^2_{diff}(2) = 6.18$, $p = .046$. Thus, sex was not further considered.

[3] Maximum-likelihood estimates of the correlations take into account missing data and nonindependence of observations. Additional correlations used in the genetic models can be obtained from Naomi P. Friedman.

Table 2
Model Fit Statistics for Reduced Phenotypic Executive Function Confirmatory Factor Analysis Models

Model	Model fit					Fit vs. full model		
	χ^2	df	p	RMSEA	TLI	χ^2_{diff}	df	p
Full three-factor model (Figure 2)	44.51	24	.007	.039	.955			
Two-factor models								
Inhibiting–Updating collapsed and Shifting	83.96	26	< .001	.063	.884	43.04	2	< .001
Inhibiting–Shifting collapsed and Updating	91.65	26	< .001	.067	.868	43.16	2	< .001
Updating–Shifting collapsed and Inhibiting	192.61	26	< .001	.106	.666	296.22	2	< .001
One-factor model	195.86	27	< .001	.105	.674	220.36	3	< .001

Note. Values of χ^2/df below 2, root-mean-square error of approximation (RMSEA) below .06, and Tucker–Lewis index (TLI) above .95 indicate good fit.
[a] χ^2 and χ^2_{diff} are scaled for nonindependence (Satorra & Bentler, 2001), so χ^2_{diff} will not equal the simple difference between χ^2s for the full and reduced models.

as well as the etiology of their unity and diversity. However, for thoroughness and to facilitate comparison to previous studies using individual tasks, we also estimated ACEs for each task separately. Table 3 presents these ACE estimates as well as the MZ and DZ correlations. As discussed earlier, MZ correlations greater than DZ correlations suggest a genetic effect, DZ correlations greater than half the MZ correlations suggest shared environmental influences, and MZ correlations less than 1.0 suggest some influence of nonshared environment (which can include measurement error at the level of individual tasks).

As shown in Table 3, the MZ correlations were uniformly higher than the DZ correlations, and in most cases, the DZ correlations were not more than half the MZ correlations, leading to estimates of moderate genetic influence (29%–56%), little to no shared environmental influence (0%–19%), and moderate nonshared environmental influence (44%–66%). These values are consistent with previous estimates of ACEs based on individual executive function measures (e.g., Ando et al., 2001; Malone & Iacono, 2002). Given the task impurity evident in the confirmatory factor

analysis (Figure 2), however, it is unclear to what extent these estimates reflect genetic and environmental influences on the underlying executive functions.

ACE Models of Individual Latent Variables: Etiology of Individual Differences in Each Executive Function

To address the first major goal of the study (specifying the etiology of the three executive functions), we applied the ACE model to each latent variable (Appendix B discusses alternative models we examined to verify the appropriateness of these latent variable models, as well as the model fit statistics). Figure 3 depicts the three separate models for Inhibiting, Updating, and Shifting. Each model includes the ACE estimates for the target executive function (e.g., A_I, C_I, and E_I for Inhibiting) as well as independent ACEs for the individual tasks, which reflect genetic and/or environmental variance that is specific to each task (i.e., not captured by the relevant latent variable).

Table 3
Univariate Task Twin Correlations and ACE Estimates

Task	Twin correlations[a]		Variance components			Model fit			
	MZ	DZ	A	C	E	$\chi^2(6)$	p	RMSEA	TLI
Antisaccade	**.56**	**.22**	**.56**	.00	**.44**	3.06	.801	.000	1.02
Stop signal	**.57**	.09	**.48**	.00	**.52**	13.86	.031	.095	.940
Stroop	**.52**	**.35**	**.38**	.14	**.48**	8.83	.183	.057	.984
Keep track	**.51**	**.18**	**.50**	.00	**.50**	2.43	.876	.000	1.03
Letter memory	**.57**	.09	**.52**	.00	**.48**	14.98	.020	.101	.949
Spatial 2-back	**.31**	.12	.29	.00	**.71**	2.67	.850	.000	1.07
Number–letter	**.49**	**.26**	**.50**	.00	**.50**	4.60	.596	.000	1.01
Color–shape	**.34**	**.19**	.29	.05	**.66**	4.07	.667	.000	1.03
Category switch	**.54**	**.38**	**.35**	.19	**.46**	4.92	.554	.000	1.01
WAIS-IQ	**.84**	**.52**	**.69**	.16	**.15**	2.46	.873	.000	1.01
Hidden patterns	**.74**	**.43**	**.69**	.07	**.25**	2.69	.847	.000	1.01
Perceptual speed	**.73**	**.24**	**.70**	.00	**.30**	5.51	.480	.000	1.00
Identical pictures	**.76**	**.37**	**.76**	.00	**.24**	2.85	.827	.000	1.01

Note. Values of χ^2/df below 2, root-mean-square error of approximation (RMSEA) below .06, and Tucker–Lewis index (TLI) above .95 indicate good fit. Boldface type indicates $p < .05$. MZ = monozygotic; DZ = dizygotic; A = additive genetic variance; C = shared environmental variance; E = nonshared environmental variance.
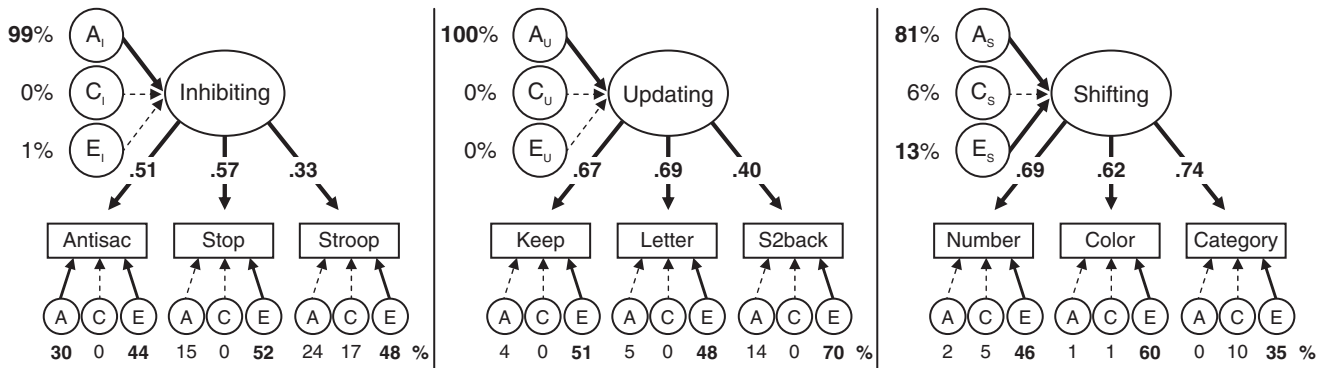
Figure 3. Three executive functions' ACE models. Numbers next to the executive functions' ACEs (e.g., A_U) are the percentages of executive function latent variable variances due to genetic (A), shared environmental (C), and nonshared environmental (E) influences. Numbers under the lower ACEs are estimates for nonexecutive variances in individual tasks. Numbers on top of the arrows from latent executive functions to tasks are standardized factor loadings. Boldface type and solid lines indicate $p < .05$. Antisac = antisaccade; stop = stop signal; keep = keep track; letter = letter memory; S2back = spatial 2-back; number = number–letter; color = color–shape; category = category switch.

As shown in Figure 3, all three executive functions showed extremely large genetic ($A_I = 99\%$, $A_U = 100\%$, and $A_S = 81\%$) and little to no significant environmental influences ($E_S = 13\%$). Hence, even though the individual tasks have only moderate genetic influences, the variance common to multiple exemplar tasks is highly heritable. These higher heritability estimates indicate that task impurity and measurement error lead to underestimates of genetic influences on executive functions that can be corrected with latent variables.

To see how the high heritability of the latent variables can arise from extracting what is common to multiple tasks with only moderate heritability, it may be helpful to see how the total variance in each task can be decomposed into the variance due to the target executive function's ACEs and the individual task's ACEs. For each model, the ACE estimates for the latent variable (e.g., A_U, C_U, and E_U for Updating) should sum to 100%, as together they account for 100% of the executive function's variance. This characteristic of ACEs summing to 100% will be the case for the variables at the highest level of each model, because nothing else predicts them. However, the ACEs for each task (e.g., the ACEs below the keep track rectangle in Figure 3) do not sum to 100%, because a particular task's total variance is a combination of both these task-specific ACEs and the ACEs for the target executive function. For the keep track task, for example, the variance due to A_U is the A_U's heritability estimate (100%) multiplied by the square of the factor loading of keep track on the Updating latent variable ($.67^2 = .45$): $100 * .45 = 45\%$. Hence, about 45% of the variance in the keep track task is due to genetic influences on Updating ability. The remaining 55% is unrelated to Updating and is shown in the ACEs specific to the keep track task (i.e., 4% A + 0% C + 51% E).

As this explanation illustrates, the genetic and environmental variance in each measure can be decomposed into that which is shared with other measures and that which is unique to each measure. This same decomposition can be applied in a model that includes multiple executive functions, as described next.

ACE Model of the Three Executive Functions Together: Etiology of the Unity and Diversity

The second major goal of the study was to specify the etiology of the unity and diversity of executive functions. Toward this goal, we estimated multivariate ACE models that included all three executive functions at once. Such models can be parameterized in a number of ways, but for the purposes of this study, two alternative parameterizations provide complementary information: (a) a hierarchical model and (b) a nested factors model. These two models are described in detail in the following sections.

Hierarchical model. The first multivariate model we estimated was the hierarchical ACE model in Figure 4, which includes all three executive function latent variables loading on a Common Executive Function (Common EF) factor. This model fit the data well, $\chi^2(322) = 384.45$, $p = .010$, RMSEA = .036, TLI = .949. As shown in Figure 4, the model includes three levels of ACEs: (a) general ACEs that influence all three executive functions through a Common EF factor (A_C, C_C, and E_C); (b) ACEs unique to each executive function (e.g., A_I, C_I, and E_I for Inhibiting); and (c) the residual ACEs for the individual tasks (i.e., the nonexecutive variance in each task).

The high heritability of the Common EF factor ($A_C = 99\%$) indicates that genetic influences mediated almost all of the variance common to the three executive functions. Using the procedures described earlier, one can decompose the proportions of variance in each executive function due to the Common EF factor's ACEs: Multiplying each executive function's squared factor loading by the A_C variance (99%) reveals that A_C explained 99%, 43%, and 44% of the variance in Inhibiting, Updating, and Shifting, respectively. Of note, the ACEs unique to each executive function indicated that Updating and Shifting also had their own significant independent genetic influences ($A_U = 56\%$ and $A_S = 42\%$) and Shifting also had small but significant nonshared environmental variance ($E_S = 13\%$). These results indicate that the
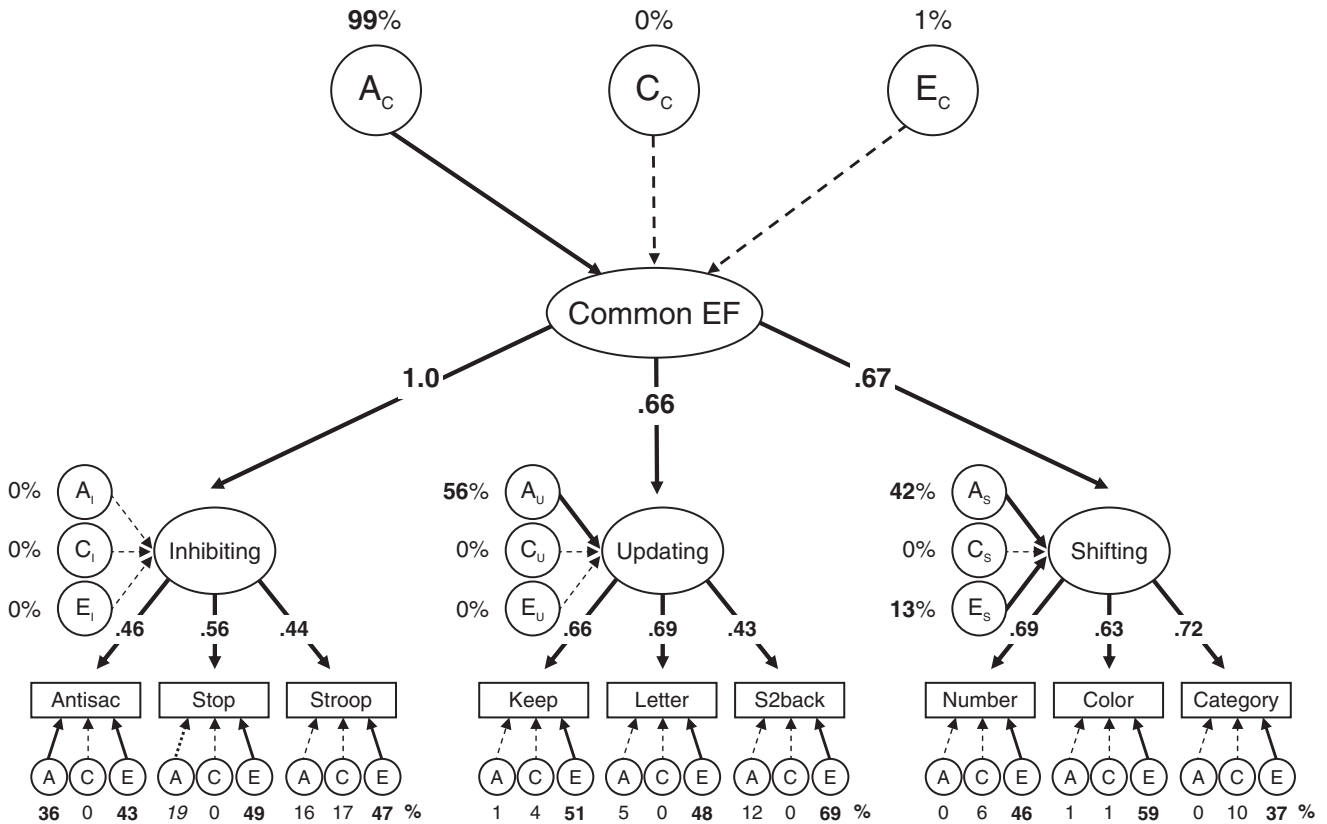
Figure 4. Hierarchical multivariate executive function ACE model. Numbers above the top ACEs (A_C, C_C, and E_C) are the percentages of the Common Executive Function (Common EF) variance due to genetic and environmental influences. Numbers occluding arrows are standardized factor loadings. Numbers next to latent variable executive functions' ACEs (e.g., A_I, A_U, and A_S) are their specific genetic and environmental variances independent of the Common EF factor. Numbers under the lower ACEs are estimates for the remaining nonexecutive variances in individual tasks. Boldface type and solid lines indicate $p < .05$. The italic type indicates stop signal's specific A variance approached significance ($p = .052$). Antisac = antisaccade; stop = stop signal; keep = keep track; letter = letter memory; S2back = spatial 2-back; number = number–letter; color = color–shape; category = category switch.

genetic structure of executive functions is fractionated. There were significant genetic influences at multiple levels: those common to all three executive functions and those specific to Updating and Shifting. Moreover, there were also some genetic influences at the third level, even though such task-specific, nonexecutive genetic effects reached statistical significance only for the antisaccade task in this model.

These findings rule out hypotheses that include environmental influences as substantial contributors to the unity and diversity of executive functions. Rather, they support the more surprising conclusion that the unity and diversity of these three executive functions are almost entirely of genetic origin: The executive functions correlate because they share common genetic influences, and they are separable mainly because of independent genetic influences on Updating and Shifting.

Nested factors model. An alternative approach to these data is the nested factors model depicted in Figure 5. The first factor in this model is the Common EF factor, on which all nine executive function tasks directly load. Though parameterized differently than the Common EF factor in Figure 4, it is conceptually the same, and

its variance directly corresponds to the Common EF factor in the hierarchical model. The nested factors are the Updating-specific factor (on which the three updating tasks load) and the Shifting-specific factor (on which the three shifting tasks load). Because the variance that is common to all nine executive function tasks (i.e., the unity) is already accounted for by the Common EF factor, these latter factors capture the remaining variance common to the updating and shifting tasks, respectively. Thus, these factors do not correspond to the Updating and Shifting factors discussed in previous models but rather can be considered components of these factors (i.e., the components that do not include what is common to all three executive functions). The significant loadings on these Updating-specific and Shifting-specific latent variables indicate that there is something common among the Updating tasks and among the Shifting tasks, over and above the Common EF factor, thus again demonstrating unity and diversity of executive functions.

In this model, there is no Inhibiting-specific factor because there is no variance in Inhibiting that is unrelated to the Common EF factor (as was the case in the hierarchical model in Figure 4, in
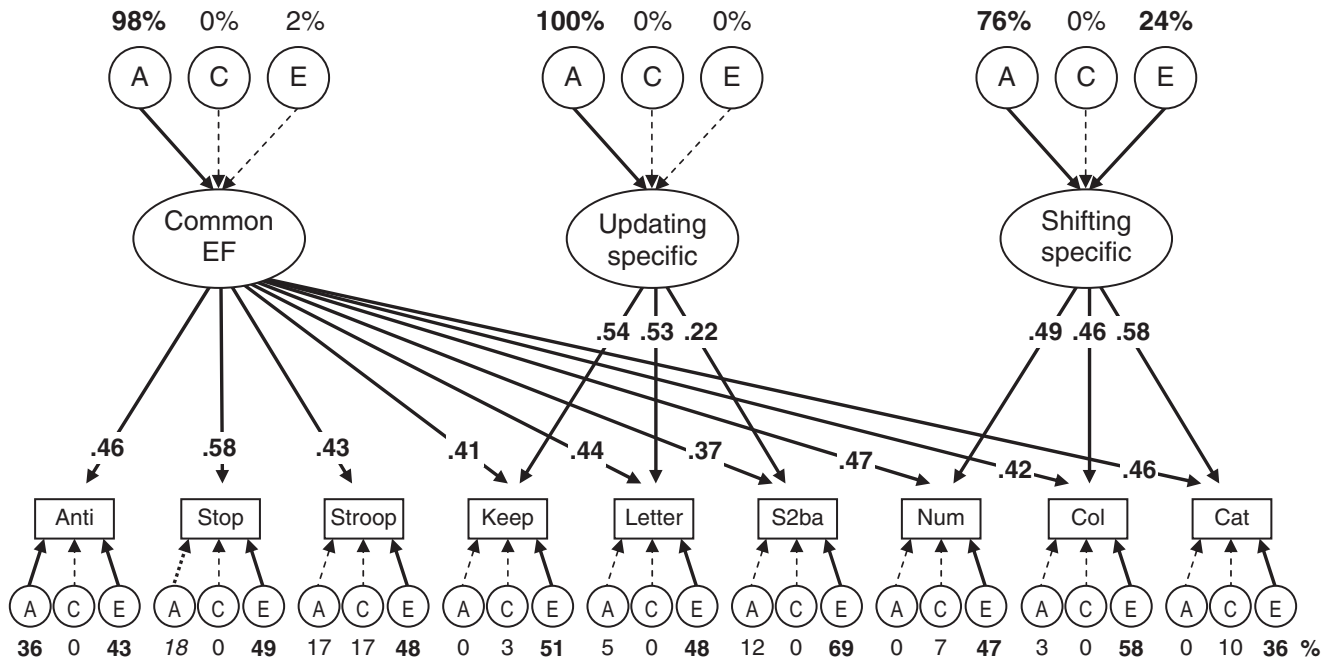
Figure 5. Nested factors multivariate executive function ACE model. All nine executive function tasks load on the first Common Executive Function (Common EF) factor; hence, this factor represents what is common to all three executive functions. The updating and shifting tasks also load on Updating-specific and Shifting-specific factors, respectively; hence, these factors represent what is specific to the updating and shifting tasks once the influence of the Common EF factor is removed. Numbers above the top ACEs are the percentages of the Common EF, Updating-specific, and Shifting-specific factors' variances due to genetic and environmental influences. Numbers occluding arrows are standardized factor loadings. Numbers under the lower ACEs are estimates for the remaining nonexecutive variances in individual tasks. Boldface type and solid lines indicate $p <$ .05. The italic type indicates stop signal's specific A variance approached significance ($p = .057$). Anti = antisaccade; stop = stop signal; keep = keep track; letter = letter memory; S2ba = spatial 2-back; num = number–letter; col = color–shape; cat = category switch.

which the Inhibiting latent variable had a loading of 1.0 on the Common EF factor.[4] It is important to note in this context that these results do not mean that there is no Inhibiting ability. As shown in the phenotypic (Figure 2) and hierarchical genetic (Figure 4) models, we were able to extract an Inhibiting latent variable that was separable from the Updating and Shifting latent variables. However, as indicated by Inhibiting's loading of 1.0 on the Common EF factor in the hierarchical genetic model (Figure 4), the variance in this latent variable was entirely explained by the variance common to all three executive functions and hence completely subsumed in the Common EF factor.

The fit of this nested factors model was similar to that of the hierarchical model, $\chi^2(321) = 378.53$, $p = .015$, RMSEA = .035, TLI = .952. As shown in Figure 5, the Updating-specific factor was entirely genetic, and the Shifting-specific factor also had significant genetic variance (as well as significant nonshared environmental variance). Hence, this model, which takes a different approach to modeling the data, shows the same pattern as the model in Figure 4.

We present both of these models because they provide complementary ways of looking at the data. The advantage of the hierarchical model is that it enables a decomposition of each latent variable (Inhibiting, Updating, and Shifting) into the variance due

to the Common EF factor's ACEs, and the ACEs specific to that executive function. The advantage of the nested factors model is that it provides three latent variables that isolate this common and specific variance at the same level (that is why the ACEs for each of the three latent variables sum to 100%). This single-level feature of the nested factors model makes it more amenable for examining genetic correlations with other variables, making it ideal for the secondary analyses presented next.

Secondary Genetic Analyses

In addition to the primary genetic analyses reported above, we examined two additional models to rule out the hypotheses that the Common EF factor simply reflects perceptual speed or *g*. The results of these analyses indicate that the Common EF factor goes beyond these two cognitive constructs.

---

[4] Allowing an Inhibiting-specific factor that loaded on the three Inhibiting tasks in the phenotypic model resulted in near-zero loadings for two of the tasks and a Heywood case (a standardized loading greater than 1.0) for the third task, indicating model misspecification. Dropping this factor did not harm model fit.

ACE Model Including Speed: Is the Genetic Variance in Executive Functions the Same as That for Speed?

We tested the hypothesis that the unity of executive functions primarily reflects speed (e.g., Salthouse, 2005) by adding a Perceptual Speed latent variable (derived from three perceptual speed tasks) to the multivariate nested factors model and allowing its A and E to correlate with the nonzero As and Es for the Common EF and nested factors, as shown in Figure 6. A correlation between the Perceptual Speed factor's A variance and Common EF factor's A variance that is significantly smaller than 1.0 would indicate that the Common EF factor's genetic variance is not completely the same as the Perceptual Speed factor's genetic variance.

This model showed good fit, $\chi^2(570) = 727.53$, $p < .001$, RMSEA = .043, TLI = .931. As shown in Figure 6, Perceptual Speed's genetic A variance significantly correlated with the Common EF factor's A variance ($r = .67$) and also with the Updating-specific factor's A variance ($r = .19$), though the latter correlation was small. The correlations between the Perceptual Speed factor's and the Shifting-specific factor's A and E variances were small and not significant (rs = .08 and .09, respectively). Of note, the genetic correlation between the Perceptual Speed factor and the Common EF factor ($r = .67$) could not be set to 1.0 without harming model fit, $\chi^2_{\text{diff}}(1) = 8.87$, $p = .003$. Also, the small correlation between the Perceptual Speed factor's A and the

Updating-specific factor's A ($r = .19$) could not be set to 1.0 without harming model fit, $\chi^2_{\text{diff}}(1) = 13.58$, $p < .001$. These findings falsify the hypothesis that the Common EF factor simply reflects processing speed, although speed is substantially related to the Common EF variance.

ACE Model With IQ: Is the Genetic Variance in the Common EF Factor the Same as That for g?

WAIS–IQ as a manifest variable. Recent neuropsychological and neuroimaging research suggests that *g* is related to lateral frontal cortex (Duncan et al., 2000), a general area that is also common to multiple executive functions (Collette et al., 2005; Sylvester et al., 2003). Moreover, *g* and IQ are known to have genetic influences (50% to 70%; Plomin & McClearn, 1993). Given that the Common EF factor captures the variance general to all three executive functions, a natural question is whether its genetic variance simply reflects the substantial genetic influences on *g*. To answer this question, we estimated the correlation between WAIS–IQ's A and E variances and the nonzero A and E variances for the three executive function latent variables in the nested factors model (Figure 7).

This model fit the data well, $\chi^2(395) = 467.47$, $p = .007$, RMSEA = .035, TLI = .956. As shown in Figure 7, the genetic variance in WAIS–IQ significantly correlated with the genetic
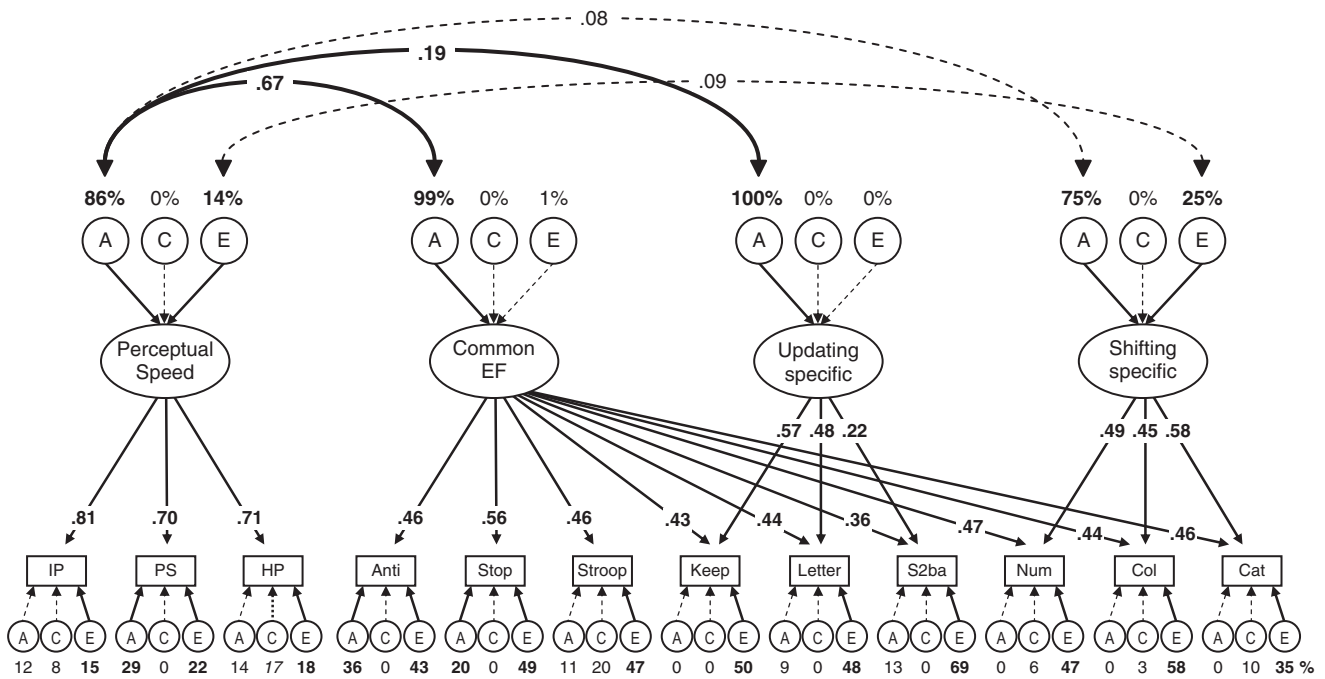
Figure 6. Executive function ACE model with Perceptual Speed. Numbers above the ACEs for the latent variables are the percentages of those latent variables accounted for by genetic and environmental influences. Numbers occluding the double-headed arrows are correlation coefficients. Correlations for components with zero or near-zero variances were not estimated. Numbers occluding arrows are standardized factor loadings. Numbers under the lower ACEs are estimates for task-specific variances. Boldface type and solid lines indicate $p < .05$. The italic type indicates hidden patterns' specific C variance approached significance ($p = .055$). IP = identical pictures; PS = perceptual speed; HP = hidden patterns; Anti = antisaccade; stop = stop signal; keep = keep track; letter = letter memory; S2ba = spatial 2-back; num = number–letter; col = color–shape; cat = category switch.
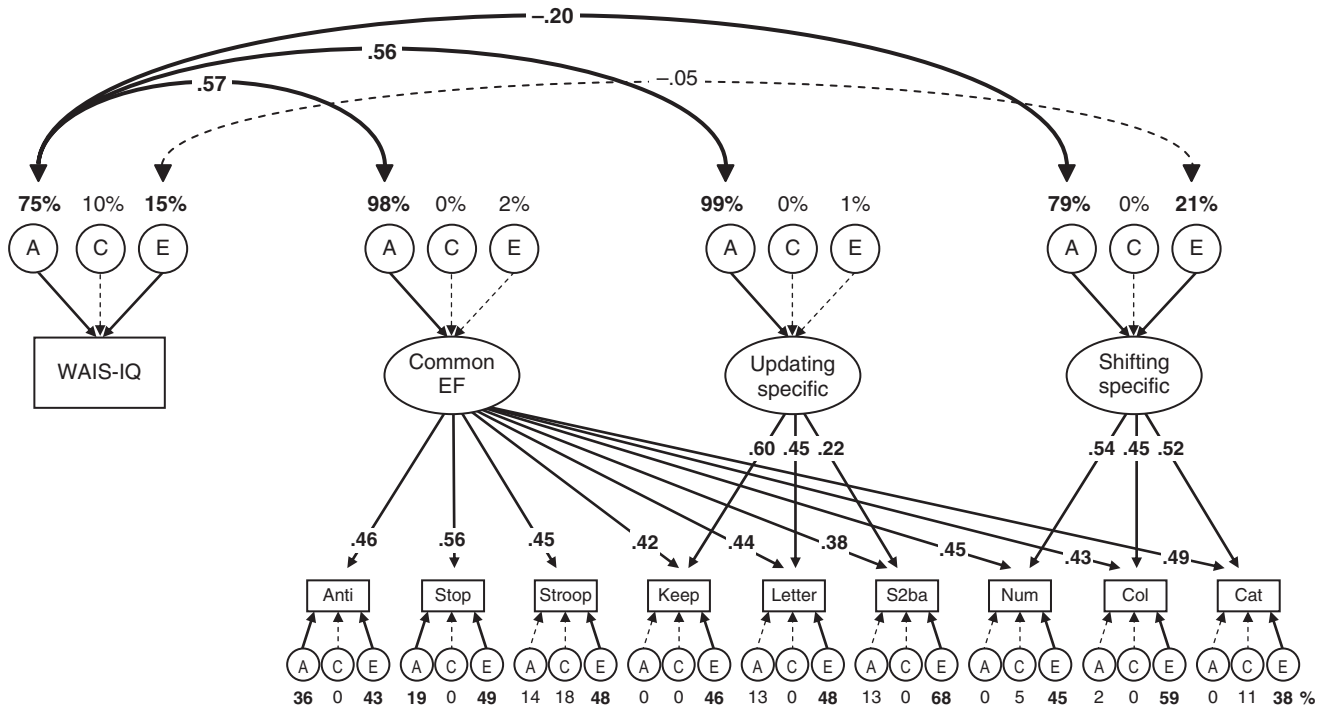
Figure 7. Nested factors executive function model with Wechsler Adult Intelligence Scale full-scale IQ (WAIS–IQ). Numbers above the ACEs for the latent variables and WAIS–IQ are the percentages of those variables accounted for by genetic and environmental influences. Numbers occluding the double-headed arrows are correlation coefficients. Correlations for components with zero or near-zero variance were not estimated. Numbers occluding arrows are standardized factor loadings. Numbers under the lower ACEs are estimates for task-specific variances. Boldface type and solid lines indicate $p < .05$. Anti = antisaccade; stop = stop signal; keep = keep track; letter = letter memory; S2ba = spatial 2-back; num = number–letter; col = color–shape; cat = category switch.

variance for the Common EF factor ($r = .57$) and also for the Updating-specific ($r = .56$) and Shifting-specific ($r = -.20$) factors. WAIS–IQ's E variance did not significantly correlate with the Shifting-specific factor's E variance. Though the correlations with the Common EF factor and the Updating-specific factor were expected on the basis of our previous phenotypic finding that IQ was more closely related to Updating than to Inhibiting and Shifting (Friedman et al., 2006), the small negative genetic correlation between IQ and the Shifting-specific factor was unexpected and difficult to interpret. Of note, the correlation between WAIS–IQ's A variance and the Common EF factor's A variance ($r = .57$) could not be set to 1.0 without harming model fit, $\chi^2_{diff}(1) = 15.03, p < .001$; nor could the correlations between WAIS–IQ's A and the As for the Updating-specific factor ($r = .56$), $\chi^2_{diff}(1) = 9.50, p = .002$, and for the Shifting-specific factor ($r = -.20$), $\chi^2_{diff}(1) = 5.90, p = .015$. Hence, these results indicate that the Common EF factor is not just g.

WAIS–IQ as a latent variable. In the analyses discussed in the previous section, g was operationalized as WAIS–III full-scale IQ, a manifest variable. One concern about this use of a single variable is that it may not be on an even footing with the executive function variables, as the latter are latent (as opposed to manifest) variables. Although the IQ scores are composites of 11 separate subtests and hence are not at the same level of impurity and measurement error

as individual tasks used to tap the three executive functions, we conducted further analyses with g as a latent variable (on which the 11 WAIS–III subtests loaded). The complexity of these analyses rendered their full presentation beyond the scope of the current article, but the main results, described next, showed essentially identical patterns as the analyses with WAIS–IQ as an observed variable.

To achieve satisfactory model fit and avoid biases in the parameters due to model misspecification, it was necessary to model not only a general g factor (our primary interest) but also four additional WAIS–III factors that have been established in previous literature (e.g., Taub, McGrew, & Witta, 2004; Wechsler, 1997). To account for these factors while still extracting a general component, we used the nested factors model depicted in Appendix C, with all 11 subtests loading on the g factor and each subtest also loading on one of four additional independent factors (with the loading pattern specified a priori on the basis of previous literature). This model provided a good fit to the data, $\chi^2(479) = 547.46, p = .016$, RMSEA = .031, TLI = .977, with all of the subtests loading significantly on the g latent variable. The g latent variable was 76% heritable, with 21% C variance and 2% E variance. Hence, even when the g variable is operationalized as a latent variable loading on 11 individual measures, its heritability is

quite similar to that of the WAIS–IQ scores, though in this case the environmental variances did not reach significance.

To examine the relationship between the g latent variable's genetic variance and the executive functions' genetic variances, we combined this nested factors g model (Appendix C) with the nested factors executive function model (Figure 5) and estimated the genetic correlations between the g latent variable and the executive function latent variables, just as we did in the models with Perceptual Speed and WAIS–IQ (Figures 6 and 7) described earlier. Note that in these models, it was necessary to include correlations with the additional independent WAIS–III factors to avoid model distortion, but we do not report these parameters here, as they are not of primary interest for the current study. This combined model provided an acceptable fit to the data, $\chi^2(1571) = 1{,}869.77$, $p < .001$, RMSEA $= .036$, TLI $= .934$.

The genetic correlations between the g latent variable and the Common EF, Updating-specific, and Shifting-specific variables were .52, .49, and −.17, respectively. These correlations are similar to those with the WAIS–IQ scores summarized in Figure 7 (.57, .56, and −.20, respectively). The correlations with the Common EF and Updating-specific (though not the Shifting-specific) latent variables were significant. Just as in the model using WAIS–IQ, however, the genetic correlations of the g factor with the Common EF factor ($r = .52$) and with the Updating-specific factor ($r = .49$) could not be constrained to 1.0 without harming model fit, $\chi^2_{\text{diff}}(1) = 17.12$, $p < .001$, and $\chi^2_{\text{diff}}(1) = 9.70$, $p = .002$, respectively. Hence, even when g is treated as a latent variable, its genetic variance does not completely explain the genetic variance common to the three executive functions.

General Discussion

Summary of the Main Results

Primary Analyses: Genetic and Environmental Influences on Executive Functions

The two primary goals of this study were to specify (a) the extent to which genetic and environmental influences lead to individual differences in three executive functions (Inhibiting, Updating, and Shifting) and (b) the extent to which the unity and diversity of executive functions are due to common and specific genetic and/or environmental influences. Our results provide clear answers to these questions.

With respect to the first goal, our results indicate that individual differences in executive functions are almost entirely genetic at the level of latent variables, placing them among the most heritable psychological traits, possibly even more heritable than IQ. The finding of virtually no environmental influences at the general (Common EF) level is particularly striking in light of the fact that moderate environmental variance has consistently been found for other cognitive functions, including specific cognitive abilities (40% to 70% environmental; Alarcón et al., 1999; Pedersen et al., 1992) and g (30% to 50% environmental; McGue et al., 1993; Neisser et al., 1996); indeed, our own univariate estimate indicated significant environmental (30%) influences for WAIS–IQ scores.

With respect to the second goal, the unity and diversity of executive functions are due almost entirely to common and unique genetic influences. Our estimates from the hierarchical model

indicated that their unity was due to the influence of a Common EF factor that was 99% heritable. Their diversity was due primarily to substantial genetic influences unique to Updating (56%) and Shifting (42%); Shifting also had small but significant (13%) nonshared environmental influences. The nested factors genetic model of the executive functions showed the same pattern. This result—that both the unity and the diversity of executive functions are primarily genetic—provides a marked contrast to findings with specific cognitive abilities (e.g., verbal and spatial abilities, memory, and speed), suggesting that they are related because of shared genetic and environmental influences and separable primarily because of environmental influences unique to particular abilities, with only small specific genetic influences (Alarcón et al., 1999; Petrill, 1997).

Secondary Analyses: Genetic Relations of Executive Functions to Speed and IQ

In two secondary analyses, we found that the genetic influences on executive functions went beyond Perceptual Speed or IQ. Specifically, although these two constructs were genetically correlated with the executive function latent variables, these correlations were far from 1.0. These results indicate that the three executive functions examined here reflect genetic variance that, though related, is not redundant with these other well-studied cognitive constructs, thus ruling out the hypothesis that once such factors are taken into account, there is not much unique variance left for executive functions (Salthouse, 2005; Salthouse et al., 2003).

The finding that Perceptual Speed was primarily related to the Common EF factor (with a small additional correlation with the Updating-specific factor) provides a new basis for interpreting recent reports that executive functions, particularly inhibition-related abilities, are related to processing speed (e.g., Hedden & Yoon, 2006; Salthouse et al., 2003). That is, speed seems to be related to all three of these executive functions because it is genetically related to what they all share (the Common EF factor), and it is slightly more related to Inhibiting because Inhibiting is more related to the Common EF factor.

The finding that the genetic variance in WAIS–IQ was substantially related to both the Common EF factor and the Updating-specific factor complements and extends our earlier phenotypic finding with these data (Friedman et al., 2006) that IQ (both its fluid and its crystallized aspects) was primarily related to Updating abilities. In particular, it seems that IQ (whether operationalized as a single manifest variable or as a latent variable) shows a stronger relationship with Updating than with the other two executive functions because it is substantially related to the genetic variance that is specific to Updating as well as the genetic variance that is common to the three executive functions. These results provide one illustration of how taking into account multiple levels of genetic variance (i.e., the common and specific genetic variance) in these three executive functions may elucidate their relations to other variables of interest.

One possible concern with these secondary analyses is that the speed and IQ measures were collected in a separate session approximately 1 year prior to the executive function session, and that this gap might have lowered their correlations with the executive function latent variables in the secondary analyses. It is important

to note, however, that our analyses were concerned only with the genetic variance in these measures. Previous research on the developmental stability and change of cognitive abilities in this age range suggests that genetic effects contribute to stability of scores, whereas instability is due primarily to nonshared environment. For example, Petrill et al. (2004) found that the genetic correlation between IQ scores at year 12 and year 16 was 1.0, strongly suggesting that there are no new sources of genetic variation in this age range (which is quite a bit larger than the range of 16 to 17 in the current study). Hence, we think it unlikely that the 1-year gap between the two sessions substantially lowered the genetic correlations between the executive functions at age 17 and the other cognitive abilities at age 16.

Provisos of Twin Studies

Twin models such as those used in this study have been criticized on a number of grounds, although some do not apply to this study (such as parents' rater bias). Perhaps the most relevant issue for our results is the possibility that the equal environments assumption (the assumption that the shared environmental influences causing similarity of twins are about the same for identical and fraternal twins) was violated. If identical twins experience more similar environments than fraternal twins, and those environments actually do influence the traits of interest, the heritability estimates may be biased upward. However, a number of studies that tested the equal environments assumption through different methodologies have found little evidence for unequal environments influencing twin similarities or differences (e.g., Kendler, Neale, Kessler, Heath, & Eaves, 1993; Loehlin & Nichols, 1976; Morris-Yates, Andrews, Howie, & Henderson, 1990; Scarr & Carter-Saltzman, 1979). Moreover, the convergence of results from twin studies with the results of other designs that do not assume equal environments (such as adoption studies, twin adoption studies, and family studies) further supports the validity of the twin design (e.g., Bouchard, Lykken, Tellegen, & McGue, 1990).

Heritability estimates may also be biased upward if the studied population is particularly homogeneous in terms of environmental factors (e.g., socioeconomic status or educational opportunities) that influence the phenotypes examined. In the extreme case, if everyone studied were exposed to exactly the same environment, there would be no environmental factors that varied, and hence, individual differences should be entirely genetic. This fact raises the question of whether the high heritability estimates in the current study might be due to an unusually homogeneous sample in terms of environmental factors. Although we cannot entirely rule out this possibility without extensive environmental data, we think that it is unlikely given the data that we do have. First, our sample is a community-wide sample, rather than a sample of individuals in a particular group (such as college students or individuals in treatment for a disorder). Second, our most relevant data (though admittedly not optimal) about the socioeconomic status of the families, year 16 twin-reported family income brackets, suggest that though the sample was somewhat skewed toward higher incomes, it was certainly not restricted to high-income families: Of the 241 families for whom these data were available, 10% fell into the <$15K bracket, 8.3% into the $15–$30K bracket, 16.2% into the $30–$45K bracket, 18.7% into the $45–$60K bracket, 13.7% into the $60–$75K bracket, and 33.2% into the

>$75K bracket. Moreover, our most relevant data (i.e., years of education) pertaining to the educational environment of the 293 families in this study at the time of their recruitment (1984 to 1989) suggest that they were reasonably representative of the population of Colorado in terms of variability: for mothers in this longitudinal sample ($M = 14.5$, $SD = 2.1$, range = 7 to 21) versus all Colorado mothers ($M = 13.0$, $SD = 2.4$); for fathers in this sample ($M = 14.7$, $SD = 2.1$, range = 7 to 21) versus all Colorado fathers ($M = 13.6$, $SD = 2.5$) (see Rhea et al., 2006, for more information about this longitudinal sample). Of course, these educational data only speak to whether the sample was homogeneous in comparison to the state. It is an open question whether these results would generalize to populations in other regions of the United States or in other countries where conditions differ.

Another consideration when interpreting the heritability estimates for the current study is that the ACE genetic models we used assume only additive genetic effects. Nonadditive effects, such as genetic dominance and epistasis, would lower the DZ correlation relative to the MZ correlation, and so our very high heritability estimates could include nonadditive genetic variance. Given that for some of the executive function tasks used in the study (e.g., keep track), DZ correlations were only about 1/4 of their MZ counterparts (see Table 2), this possibility cannot be entirely ruled out.[5] However, when we did test for genetic dominance variation, it was not significant. It is also important to emphasize that our main goal in this study was to uncover the multivariate genetic and environmental structure underlying executive functions, rather than to partition the genetic effects into additive and dominance deviations. Thus, even if some of the executive functions' genetic variation in fact turns out to be nonadditive, our conclusions about the multilevel genetic structure of executive functions will still hold.

A final concern with twin studies in general is that twins may not be representative of the general population. Perhaps because twins share a womb and are exactly the same age, they may experience more shared environment than their nontwin siblings, leading to higher estimates of shared environmental influences compared with studies using nontwin designs, such as adoption studies (Chipuer et al., 1990). Contrary to this prediction, there was no evidence for shared environmental influences (except for IQ) in the current study. Another reason that twins may not be representative of the general population is that they are often born premature and experience different prenatal and possibly postnatal environments than singletons. However, studies comparing twins with their nontwin relatives on a variety of measures have typically found only very small differences in psychiatric measures (e.g., Kendler, Martin, Heath, & Eaves, 1995; Rhea, Corley, & Millikan, 2002) and no significant differences in IQ (Posthuma, de Geus,

---

[5] DZ correlations of only about 1/4 MZ correlations can indicate genetic dominance (D). With MZ and DZ twins reared together, it is not possible to estimate the effects of both shared environment (C) and dominant genetic (D) influences, although one or the other source can be assumed to be absent depending on whether the DZ twin correlation is greater (D = 0) or less (C = 0) than half the MZ correlation. Because in the current study the MZ correlations appeared greater than twice the DZ correlations for only some of the measures, we used the ACE models throughout. Moreover, when we estimated ADE models, the D parameters were nonsignificant.

Bleichrodt, & Boomsma, 2000). In fact, the mean IQ of the current sample (102) was roughly equivalent to that of the general population (mean of 100), and the factor structure of executive functions replicated that found by Miyake et al. (2000) with college students. These results make it likely that the current findings will generalize to a general population, nontwin sample.

Interpretation of High Heritability for Executive Functions

When confronted with an estimate of extremely high heritability of executive functions, one's first reaction may be despair that executive control abilities are fixed and immutable (one might also think that because Shifting was the only executive function to show significant environmental influences, it is the only one amenable to change). In other words, one might think that our results necessitate the interpretation that executive control abilities are all "nature" and no "nurture" and thereby fundamentally challenge recent research on environmental manipulations to improve executive function performance. Such interpretations are incorrect, however. Contrary to common misconceptions, high heritability does not mean that environmental factors cannot and do not affect executive functions.

In fact, our results do not negate the findings of recent studies demonstrating that it is possible to improve executive control abilities through targeted training. Although many training studies have focused on rehabilitation of executive functions in clinical populations (e.g., Klingberg et al., 2005; Sammer, Reuter, Hullmann, Kaps, & Vaitl, 2006; White & Shah, 2006) or in aging or developmental populations (Dowsett & Livesey, 2000; Kramer, Larish, & Strayer, 1995; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005), some research has also demonstrated improvement in normal young adults (e.g., Erickson et al., 2007; Kramer et al., 1995; White & Shah, 2006). Though the extent to which such training effects transfer to other untrained tasks or other executive functions is unclear at this point, the existing evidence clearly suggests that it is possible to improve one's performance on specific executive tasks through targeted training.

How can these results be reconciled with our finding of high heritability for executive functions? Generally speaking, heritability is an estimate of the genetic influence on individual differences around a population mean (i.e., the population variance), rather than an estimate of the influences on the mean itself. Thus, environmental factors can influence a population's average at the same time that genetic factors influence its variance (Scarr, 1992). The effects of training on executive functions may be analogous. Targeted training can certainly improve individuals' performances on executive function tasks (hence affecting the group mean), but as long as training effects influence all individuals roughly equally (reflecting an equal change from baseline across individuals), individual differences in executive functions could still be due almost entirely to genetic influences. Moreover, it is important to note that differences in environments can themselves be genetically mediated (e.g., individuals predisposed to have good executive functions might select environments that nurture those executive functions), leading to gene–environment correlations (Neale & Cardon, 1992). In such cases, environmental effects are included in the estimate for heritability, because they are themselves genetically influenced. These considerations clearly suggest that the current findings should not be taken as evidence that executive control abilities are immutable and cannot be influenced by environmental manipulations.

Implications for Future Executive Function Research

The finding that executive functions have large genetic influences at the level of latent variables is good news for researchers interested in using genetic methods to uncover the biological mechanisms underlying executive control. The genetic influences on executive functions have been of increasing interest to psychologists in a wide variety of domains (e.g., Diamond, Briand, Fossella, & Gehlbach, 2004; Goldberg et al., 2003; Nigg, Blaskey, Stawicki, & Sachek, 2004). In particular, the findings that weaknesses in executive functions are associated with numerous psychological and behavioral problems have led many researchers to suggest that executive control may be an endophenotype—a putatively more elementary characteristic closer to the underlying biology that bridges the gap between genes and disease—for these disorders. The idea behind endophenotypes is that they may provide cleaner, more reliable measures that may consequently show larger genetic effect sizes than the associated diseases (e.g., schizophrenia or ADHD). Hence, understanding the genetics of endophenotypes may be more tractable than the genetics of the diseases themselves.

Although executive functions satisfy many criteria for endophenotypes (such as being heritable, related to problems of interest, and lower in unaffected family members of individuals affected with a disorder vs. the general population; e.g., Goldberg et al., 2003), the issue of executive functions as endophenotypes has been a topic of considerable debate (see Flint & Munafò, 2006, for a recent discussion). In their meta-analyses, Flint and Munafò found little evidence that effect sizes for particular genetic loci such as the COMT Val/Met polymorphism were larger for executive function tasks (such as the Wisconsin Card Sorting Test and *n*-back) than those for the diseases of interest, and they cautioned that endophenotypes may not necessarily demonstrate simpler genetic architecture. Whether or not executive functions can truly be considered endophenotypes of certain psychiatric disorders, our findings suggest a way to increase the power and impact of genetic studies of executive functions. In particular, our results suggest that genetic sources of individual differences in executive functions may be more powerfully and purely assessed at the level of latent variables.

Another major finding of this study—that genetic influences on executive functions operate at multiple levels—is in some ways not-so-good news for researchers interested in the neurobiology of executive functions. The refutation of the hypothesis that the genetic influences on executive functions operate only at a general (Common EF) level poses a problem for many computational modeling, heritability, and molecular genetic studies that implicitly make this assumption. For example, molecular genetic studies often focus on single complex measures like the Wisconsin Card Sorting Test to examine the influences of candidate genes on executive functions (see Goldberg & Weinberger, 2004). If it were indeed the case that the genetic influences on one executive function task were general to all executive functions and that multiple executive functions were differentiated only by environmental factors, the interpretation of results using single tasks would be relatively unproblematic. However, the finding that executive functions are influenced by both general and specific genetic factors complicates the interpretation: At which level are

observed effects operating? The interpretation is made even more difficult by the fact that some of the individual tasks also showed genetic influences on the nonexecutive variance (e.g., 36% for antisaccade in the multivariate model, as shown in Figure 4). These added levels of complexity mean that progress in specifying the genetic underpinnings of executive functions will require the consideration of multiple executive functions, as well as multiple tasks per executive function, to distinguish these general and specific influences. Although doing so would no doubt be more challenging, the results of the current study suggest that this approach will likely lead to a better and more accurate understanding of the nature of genetic influences on executive control abilities.

Implications of the Lack of Inhibiting-Specific Variance

As evidenced by its loading of 1.0 on the Common EF factor in the hierarchical model (Figure 4), the variance for the Inhibiting latent variable was estimated to be shared completely with the common variance in the other two executive functions. Given that this factor loading had a confidence interval of .92 to 1.0, it is possible that Inhibiting does have some independent variance, but such variance would have to be quite small. This result was corroborated by our failure to find an Inhibiting-specific factor in the nested factors genetic model (Figure 5). Of interest, our finding that Inhibiting did not have genetic influences beyond those on the Common EF factor is reminiscent of Collette et al.'s (2005) neuroimaging finding that beyond the frontal regions activated by all three executive functions, there were no other brain regions uniquely activated by inhibiting tasks (unless statistical thresholds were lowered), even though there were such unique regions for updating and shifting tasks. That is not to say that there is no inhibiting ability. In our analyses, we did find that the inhibiting tasks clustered together to form a latent variable (replicating our earlier finding in an independent data set; Miyake et al., 2000); Collette et al. also found areas common to their inhibiting tasks. It is just to say that individual differences in response inhibition abilities appear to be very closely related to what is common among executive functions.

What does this lack of unique variance for the Inhibiting latent variable mean? One way to interpret this result may be in terms of theories that postulate inhibition as a fundamental, unifying component of executive control from which other executive deficits stem (e.g., Zacks & Hasher, 1994). Although this interpretation is consistent with the finding, one weakness of this view is that in such contexts, *inhibition* is conceptualized more broadly, including a number of possibly separable inhibition-related abilities (Friedman & Miyake, 2004; Harnishfeger, 1995; Nigg, 2000). More specifically, for this interpretation to be correct, one would have to assume that one's ability to override prepotent responses (inhibiting) is the same as the ability to remove no longer relevant information from working memory (associated with updating) as well as the ability to abandon or inhibit no longer relevant task sets (associated with shifting). In light of recent evidence that some inhibition-related control abilities (such as prepotent response inhibition and resisting proactive interference) are correlationally dissociable (Friedman & Miyake, 2004; Hedden & Yoon, 2006), such an assumption may not be tenable.

An alternative interpretation of Inhibiting's high loading on the Common EF factor (and the one we favor) is that the inhibition of prepotent responses perhaps depends particularly heavily on other processes that are fundamental to all three executive functions. One such process may be the active maintenance and management of task goals in the face of interference (Miyake et al., 2000; O'Reilly, Braver, & Cohen, 1999). Indeed, the frontal–parietal brain network activated by multiple executive functions is thought to subserve goal maintenance and selective attention or interference control (Collette et al., 2005; Sylvester et al., 2003). From this perspective, the key ability underlying the Inhibiting factor is the active maintenance and management of the current task goals, and once this ability is taken into account, not much is left to explain for the Inhibiting factor, whereas the Updating and Shifting factors involve additional abilities that go beyond active goal maintenance and management (e.g., the monitoring of the current contents of working memory for Updating). Although this alternative explanation is speculative and needs further systematic research, it has the advantage of being consistent with recent goal-directed views of inhibitory control (e.g., Kane & Engle, 2003; Morton & Munakata, 2002).

In our own research, including the current study, Inhibiting measures have consistently shown the lowest within-construct correlations (Friedman & Miyake, 2004; Miyake et al., 2000). Such low correlations have led some researchers to be unable to construct an inhibiting latent variable (e.g., Huizinga et al., 2006; van der Sluis et al., 2007), though those studies were developmental rather than focused on adult-level executive functions. Yet in the current study we found Inhibiting to be the construct most closely related to the Common EF factor. In other research using these data, we have consistently found the Inhibiting construct to be more closely related than Updating or Shifting to various real-world problems, such as attention problems (Friedman et al., 2007), depressive symptoms (Sabella et al., 2007), and externalizing behavior (Young et al., 2007). Others have also found prepotent response inhibition (usually measured with the stop-signal task) to be more related than other executive functions to some forms of psychopathology, such as ADHD and substance use disorders (e.g., Nigg et al., 2006; Willcutt et al., 2001). Hence, although the inhibition of dominant responses is a difficult construct to measure (see Friedman & Miyake, 2004, for discussion), it seems to be a crucially important one when it comes to understanding executive control and its relation to real-world behavior.

Conclusions and Future Directions

The results of the current study suggest that individual differences in normal young adults' Inhibiting, Updating, and Shifting abilities are almost entirely genetic in origin and that the unity and diversity of these executive functions are due primarily to genetic influences that operate at both the general level (on all three executive functions) and specific levels (on Updating and Shifting). Our finding that executive functions have both common and specific genetic influences suggests that an important goal for future genetic research will be to begin the search for specific genetic polymorphisms associated with each level. Doing so will provide information instrumental to testing and extending current theories, yielding new insights into the nature of executive control and various disorders linked with executive deficits.

In this regard, one important direction for such research may be to integrate brain localization data with emerging information

about the neuropsychological consequences of various key genetic polymorphisms. As noted earlier, the multilevel genetic structure of executive functions is consistent with recent findings of unity and diversity of executive function localization (Collette et al., 2005; Sylvester et al., 2003). This correspondence between behavioral genetic and brain localization results suggests that genetic regulation of brain structure differences (Thompson et al., 2001) and differential gene expression in these common and unique brain areas may explain some of the genetic unity and diversity of executive functions, thus making the exploration of this potential relationship a highly promising research strategy in this field.

Another promising direction for future genetic research, which has already received a good deal of attention, is to specify the role of genes that regulate neurotransmitter systems in modulating executive functions. In particular, although dopamine has long been thought to enhance flexible information updating and reduce distraction by optimizing signal-to-noise ratio in the frontal cortex, the specific mechanisms of dopamine neuromodulation are not yet fully understood (Cohen, Braver, & Brown, 2002). Recent molecular genetic studies of normal adults have found some evidence that genes influencing dopamine production and regulation, such as the COMT Val/Met polymorphism, are related to executive control ability (Goldberg & Weinberger, 2004). However, these studies have predominantly limited their focus to only a single or a small subset of dopamine-related genes and examined their relations to executive control with a single complex executive function task such as the Wisconsin Card Sorting Test, thus making it impossible to distinguish multiple levels of genetic influences on executive functions. Thus, more complex designs (e.g., examining multiple executive functions using multiple exemplar tasks for each executive function) will be needed for gene identification to be truly effective.

In conclusion, the current research has demonstrated that the unity and diversity of executive functions—a pattern that has been shown repeatedly in prior phenotypic studies of executive functions (e.g., Duncan et al., 1997; Friedman et al., 2006; Hedden & Yoon, 2006; Miyake et al., 2000; van der Sluis et al., 2007)—are almost entirely genetic in origin, at least among young adults. This is a striking and surprising finding that no existing theories of executive control would have predicted a priori. By uncovering the multilevel genetic structure of executive functions for the first time, the current study thus lays the necessary groundwork for improving the precision and impact of future genetic and neuroscience research into the mechanisms underlying executive control.

References

Alarcón, M., Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1999). Molarity not modularity: Multivariate genetic analysis of specific cognitive abilities in 16-year-old children in the Colorado Adoption Project. *Cognitive Development, 14,* 175–193.

Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavior Genetics, 31,* 615–624.

Anokhin, A. P., Heath, A. C., & Myers, E. (2004). Genetics, prefrontal cortex, and cognitive control: A twin study of event-related brain potentials in a response inhibition task. *Neuroscience Letters, 368,* 314–318.

Anokhin, A. P., Heath, A. C., & Ralano, A. (2003). Genetic influences on frontal brain function: WCST performance in twins. *NeuroReport, 14,* 1975–1978.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A,* 5–28.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive function: Constructing a unified theory of ADHD. *Psychological Bulletin, 121,* 65–94.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Bouchard, T. J., Jr., Lykken, D. T., Tellegen, A., & McGue, M. (1990, October 12). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science, 250,* 223–228.

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

Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models.* New York: Springer-Verlag.

Chipuer, H. M., Rovine, M., & Plomin, R. (1990). LISREL modeling: Genetic and environmental influences on IQ revisited. *Intelligence, 14,* 11–29.

Cohen, J. D., Braver, T. S., & Brown, J. W. (2002). Computational perspectives on dopamine function in the prefrontal cortex. *Current Opinion in Neurobiology, 12,* 223–229.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, & Computers, 25,* 257–271.

Collette, F., Van Der, L. M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping, 25,* 409–423.

DeFries, J. C., Plomin, R., Vandenberg, S. G., & Kuse, A. R. (1981). Parent–offspring resemblance for cognitive abilities in the Colorado Adoption Project: Biological, adoptive, and control parents and one-year-old children. *Intelligence, 5,* 245–277.

Diamond, A., Briand, L., Fossella, J., & Gehlbach, L. (2004). Genetic and neurochemical modulation of prefrontal cognitive functions in children. *American Journal of Psychiatry, 161,* 125–132.

Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology, 36,* 161–174.

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

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000, July 21). A neural basis for general intelligence. *Science, 289,* 457–460.

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., et al. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences, USA, 98,* 6917–6922.

Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Erickson, K. I., Colcombe, S. J., Wadhwa, R., Bherer, L., Peterson, M. S., Scalf, P. E., et al. (2007). Training-induced functional activation changes in dual-task processing: An fMRI study. *Cerebral Cortex, 17,* 192–204.

Fan, J., Wu, Y., Fossella, J. A., & Posner, M. I. (2001). Assessing the heritability of attentional networks. *BMC Neuroscience, 2,* 14.

Faraone, S. V., Perlis, R. H., Doyle, A. E., Smoller, J. W., Goralnick, J. J., Holmgren, M. A., & Sklar, P. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57,* 1313–1323.

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

Flint, J., & Munafò, M. R. (2006). The endophenotype concept in psychiatric genetics. *Psychological Medicine, 37,* 163–180.

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A., Young, S. E., Corley, R. P., & Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functions in late adolescence. *Psychological Science, 18,* 893–900.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133,* 101–135.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17,* 172–179.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia.* Hove, England: Erlbaum.

Garavan, H., & Stout, J. C. (2005). Neurocognitive insights into substance abuse. *Trends in Cognitive Sciences, 9,* 195–201.

Geurts, H. M., Verté, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry, 45,* 836–854.

Goldberg, T. E., Egan, M. F., Gscheidle, T., Coppola, R., Weickert, T., Kolachana, B. S., et al. (2003). Executive subprocesses in working memory: Relationship to Catechol-O-methyltransferase Val158Met genotype and schizophrenia. *Archives of General Psychiatry, 60,* 889–896.

Goldberg, T. E., & Weinberger, D. R. (2004). Genes and the parsing of cognitive processes. *Trends in Cognitive Sciences, 8,* 325–335.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). San Diego, CA: Academic Press.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention & performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). Cambridge, MA: MIT Press.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. C., & Curtis, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded.* Odessa, FL: Psychological Assessment Resources.

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20,* 511–528.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424–453.

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

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70.

Kendler, K. S., Martin, N. G., Heath, A. C., & Eaves, L. J. (1995). Self-report psychiatric symptoms in twins and their nontwin relatives: Are twins different? *American Journal of Medical Genetics, 60,* 588–591.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A test of the equal-environments assumption in twin studies of psychiatric illness. *Behavior Genetics, 23,* 21–27.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford Press.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., et al. (2005). Computerized training of working memory in children with ADHD: A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44,* 177–186.

Kramer, A. F., Larish, J. F., & Strayer, D. L. (1995). Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied, 1,* 50–76.

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

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Loehlin, J. C., & Nichols, J. (1976). *Heredity, environment, and personality.* Austin, TX: University of Texas Press.

Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Logie, R. H., Cocchini, G., Della Sala, S., & Baddeley, A. D. (2004). Is there a specific executive capacity for dual task coordination? Evidence from Alzheimer's disease. *Neuropsychology, 18,* 504–513.

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

Luciano, M., Wright, M., Smith, G. A., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2001). Genetic covariance among measures of information processing speed, working memory, and IQ. *Behavior Genetics, 31,* 581–592.

Lyon, G. R., & Krasnegor, N. A. (Eds.). (1996). *Attention, memory, and executive function.* Baltimore: Brookes Publishing.

Malone, S. M., & Iacono, W. G. (2002). Error rate or the antisaccade task: Heritability and developmental change in performance among preadolescent and late-adolescent female twin youth. *Psychophysiology, 39,* 664–673.

Marsh, H. W., Balla, J., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. In G. S. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 315–353). Mahwah, NJ: Erlbaum.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1124–1140.

McGue, M., Bouchard, T. J., Jr., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive abilities: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 59–76). Washington, DC: American Psychological Association.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J.-C. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115,* 123–142.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (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.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81,* 111–121.

Morris-Yates, A., Andrews, G., Howie, P., & Henderson, S. (1990). Twins: A test of the equal environments assumption. *Acta Psychiatrica Scandinavica, 81,* 322–326.

Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology, 40,* 255–265.

Muthén, L. K., & Muthén, B. O. (2006). *Mplus user's guide* (4th ed.). Los Angeles: Muthén & Muthén.

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

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families.* Dordrecht, the Netherlands: Kluwer Academic.

Nichols, R. C., & Bilbro, W. C., Jr. (1966). The diagnosis of twin zygosity. *Acta Geneticae Medicae et Gemellologiae, 16,* 265–275.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77–101.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126,* 220–246.

Nigg, J. T. (2006). *What causes ADHD? Understanding what goes wrong and why.* New York: Guildford Press.

Nigg, J. T., Blaskey, L. G., Stawicki, J. A., & Sachek, J. (2004). Evaluating the endophenotype model of ADHD neuropsychological deficit: Results for parents and siblings of children with ADHD combined and inattentive subtypes. *Journal of Abnormal Psychology, 113,* 614–625.

Nigg, J. T., Wong, M. M., Martel, M. M., Jester, J. M., Puttler, L. I., Glass, J. M., et al. (2006). Poor response inhibition as a predictor of problem drinking and illicit drug use in adolescents at risk for alcoholism and other substance use disorders. *Journal of the Academy of Child and Adolescent Psychiatry, 45,* 468–475.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 375–411). New York: Cambridge University Press.

Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3,* 346–353.

Petrill, S. A. (1997). Molarity versus modularity of cognitive functioning? A behavioral genetic perspective. *Current Directions in Psychological Science, 4,* 96–99.

Petrill, S. A., Lipton, P. A., Hewitt, J. K., Plomin, R., Cherny, S. S., Corley, R., & DeFries, J. C. (2004). Genetic and environmental contributions to general cognitive ability through the first 16 years of life. *Developmental Psychology, 40,* 805–812.

Phillips, L. H. (1997). Do "frontal tests" measure executive function? Issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 191–213). Hove, England: Psychology Press.

Plomin, R., & McClearn, G. E. (Eds.). (1993). *Nature, nurture, and psychology.* Washington, DC: American Psychological Association.

Posthuma, D., de Geus, E. J. C., Bleichrodt, N., & Boomsma, D. I. (2000). Twin-singleton differences in intelligence? *Twin Research, 3,* 83–87.

Posthuma, D., Mulder, E. J., Boomsma, D. I., & de Geus, E. J. (2002). Genetic analysis of IQ, processing speed and stimulus–response incongruency effects. *Biological Psychology, 61,* 157–182.

Rhea, S. A., Corley, R. P., & Millikan, E. (2002). Testing the representativeness of adolescent twins: Data from the Colorado longitudinal and community twin studies. *Behavior Genetics, 32,* 482–483.

Rhea, S. A., Gross, A. A., Haberstick, B. C., & Corley, R. P. (2006). Colorado Twin Registry. *Twin Research and Human Genetics, 9,* 941–949.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123,* 374–393.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences, USA, 102,* 14931–14936.

Russell, J. (1997). *Autism as an executive disorder.* Oxford, England: Oxford University Press.

Sabella, S. A., Miyake, A., Friedman, N. P., Young, S. E., & Hewitt, J. K. (2007). *Not all executive functions are equally related to depression.* Manuscript submitted for publication.

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19,* 532–545.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132,* 566–594.

Sammer, G., Reuter, I., Hullmann, K., Kaps, M., & Vaitl, D. (2006). Training of executive functions in Parkinson's disease. *Journal of the Neurological Sciences, 248,* 115–119.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66,* 507–514.

Scarr, S. (1992). Developmental theories for the 1990s: Developmental and individual differences. *Child Development, 63,* 1–19.

Scarr, S., & Carter-Saltzman, L. (1979). Twin method: Defense of a critical assumption. *Behavior Genetics, 9,* 527–542.

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

Süβ, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory explains reasoning ability—and a little bit more. *Intelligence, 30,* 261–288.

Sylvester, C. Y., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., & Jonides, J. (2003). Switching attention and resolving interference: FMRI measures of executive functions. *Neuropsychologia, 41,* 357–370.

Taub, G. E., McGrew, K. S., & Witta, E. L. (2004). A confirmatory analysis of the factor structure and cross-age invariance of the Wechsler Adult Intelligence Scale—Third Edition. *Psychological Assessment, 16,* 85–89.

Teuber, H.-L. (1972). Unity and diversity of frontal lobe functions. *Acta Neurobiologiae Experimentalis, 32,* 615–656.

Thompson, P. M., Cannon, T. D., Narr, K. L., van Erp, T., Poutanen, V.-P., Huttunen, M., et al. (2001). Genetic influences on brain structure. *Nature Neuroscience, 4,* 1253–1258.

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

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corporation.

White, H. A., & Shah, P. (2006). Training attention-switching ability in adults with ADHD. *Journal of Attention Disorders, 10,* 44–53.

Wilcox, R. R., & Keselman, H. J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods, 8,* 254–274.

Willcutt, E. G., Pennington, B. F., Boada, R., Tunick, R. A., Ogline, J., Chhabildas, N. A., & Olson, R. K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 110,* 157–172.

Wright, M., de Geus, E., Ando, J., Luciano, M., Posthuma, D., Ono, Y., et al. (2001). Genetics of cognition: Outline of a collaborative twin study. *Twin Research, 4,* 48–56.

Yehene, E., & Meiran, N. (2007). Is there a general task switching ability? *Acta Psychologica, 126,* 169–195.

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5,* 7–17.

Young, S. E., Friedman, N. P., Miyake, A., Willcutt, E. G., Corley, R. P., Haberstick, B. C., & Hewitt, J. K. (2007). *Behavioral disinhibition: Its structure and relation to response inhibition across adolescence.* Manuscript submitted for publication.

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 241–264). San Diego, CA: Academic Press.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1,* 198–226.

Appendix A

Phenotypic Correlations Among the Individual Tasks

Task	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Antisaccade	—												
2. Stop signal	.31	—											
3. Stroop	.16	.18	—										
4. Keep track	.17	.27	.22	—									
5. Letter memory	.25	.19	.25	.45	—								
6. Spatial 2-back	.22	.28	.12	.29	.28	—							
7. Number–letter	.16	.28	.23	.15	.23	.15	—						
8. Color–shape	.21	.25	.27	.17	.18	.14	.40	—					
9. Category switch	.19	.30	.24	.17	.17	.17	.50	.48	—				
10. WAIS–IQ	.25	.25	.27	.52	.40	.31	.07	.14	.19	—			
11. Hidden patterns	.23	.23	.31	.27	.22	.18	.26	.24	.28	.47	—		
12. Perceptual speed	.17	.27	.34	.36	.35	.24	.19	.20	.16	.44	.46	—	
13. Identical pictures	.22	.23	.21	.27	.23	.18	.27	.27	.26	.42	.59	.57	—

Note. Correlations are maximum-likelihood estimates (from Mplus) based on all data, adjusted for missingness and nonindependence. Total $N = 582$. Directionality of the reaction time measures was reversed so that for all tasks, higher scores indicate better performance. WAIS–IQ = Wechsler Adult Intelligence Scale full-scale IQ.

Appendix B

Common Pathway Versus Independent Pathways Model Comparisons

The structures of the genetic models we present were motivated by our latent variable approach. They include assumptions about the multivariate genetic structure: in particular, that the executive functions' ACE variables influence the individual executive function tasks through their respective latent variables (i.e., Inhibiting, Updating, and Shifting). This type of model is known as a *common pathway* model (Neale & Cardon, 1992). However, there is another, less restrictive *independent pathways* model in which the common ACEs influence each task separately, not through a latent variable. These two models are depicted conceptually in Figure B1 (Panel A for the common pathway model and Panel B for the independent pathways model). The primary difference is that in the independent pathways model, there are more free parameters (three paths from each common A, C, and E), and the variance in

each task due to the common ACEs is not necessarily proportional to its relation to a latent variable. In the common pathway model, there is only one path from each executive function A, C, and E, and the variance in each task explained by these executive function ACEs is proportional to the task's loading on the latent variable.

Table B1 presents the model fits for the independent pathways and common pathway models applied to the three executive functions individually (top three rows of Table B1) as well as to the hierarchical model with all three executive functions (bottom row of Table B1; in that case the common pathway was through the Common EF factor). In all cases, the common pathway models did not fit significantly worse than the independent pathways models, confirming our hypothesis that the higher level ACEs influenced the measures through the theoretically defined latent variables.
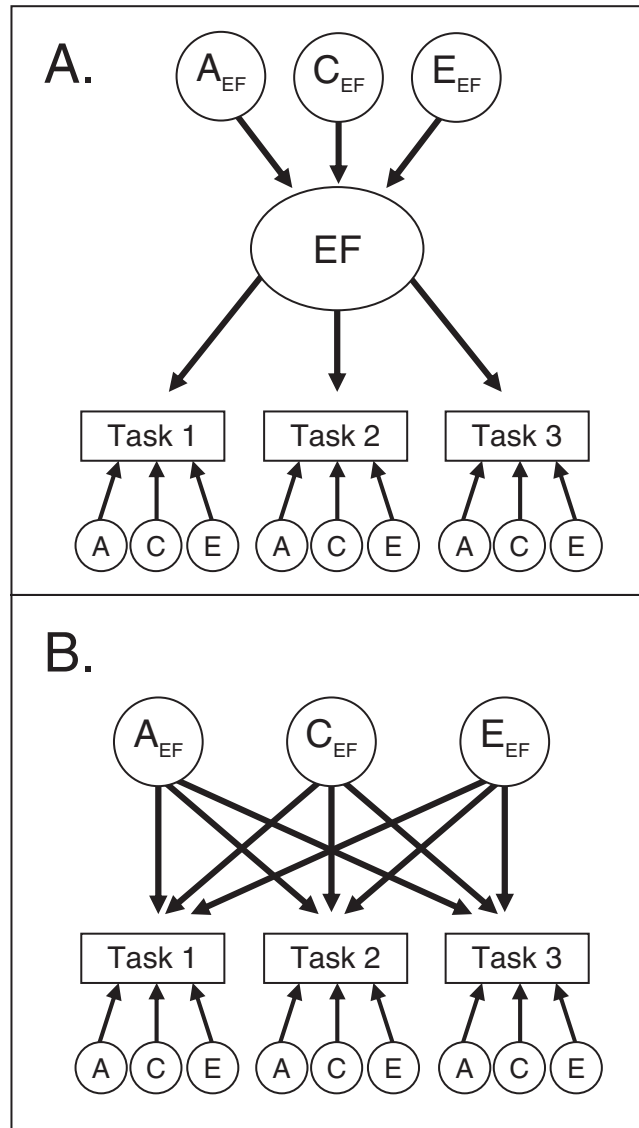
(*Appendixes continue*)

Figure B1. Two multivariate ACE models for three tasks. In a common pathway model (Panel A), the executive functions' ACE variables influence the individual tasks through the executive function (EF) latent variable. In the less restrictive independent pathways model (Panel B), the executive functions' ACEs influence the tasks directly through separate paths (i.e., there is no underlying latent variable postulated for the correlations among the measure).

Table B1
*Fit Statistics for the Common Pathway Versus Independent Pathways Models*

Common factor	Common pathway					Independent pathways					χ^2 difference		
	χ^2	df	p	RMSEA	TLI	χ^2	df	p	RMSEA	TLI	χ^2_{diff}	df	p
Inhibiting	50.90	37	.064	.051	.951	47.47	33	.049	.055	.943	3.43	4	.489
Updating	34.56	37	.584	.000	1.01	33.38	33	.449	.009	.999	1.18	4	.881
Shifting	45.86	37	.151	.040	.983	40.31	33	.178	.039	.985	5.55	4	.235
All three EFs	384.45	322	.010	.036	.949	380.94	318	.009	.037	.947	3.51	4	.476

Note. Values of χ^2/df below 2, root-mean-square error of approximation (RMSEA) below .06, and Tucker–Lewis index (TLI) above .95 indicate good fit. Nonsignificant chi-square differences indicate that the reduced common pathway models did not provide a significantly worse fit than the independent pathways models. EF = executive function.

Appendix C

g as a Latent Variable Model
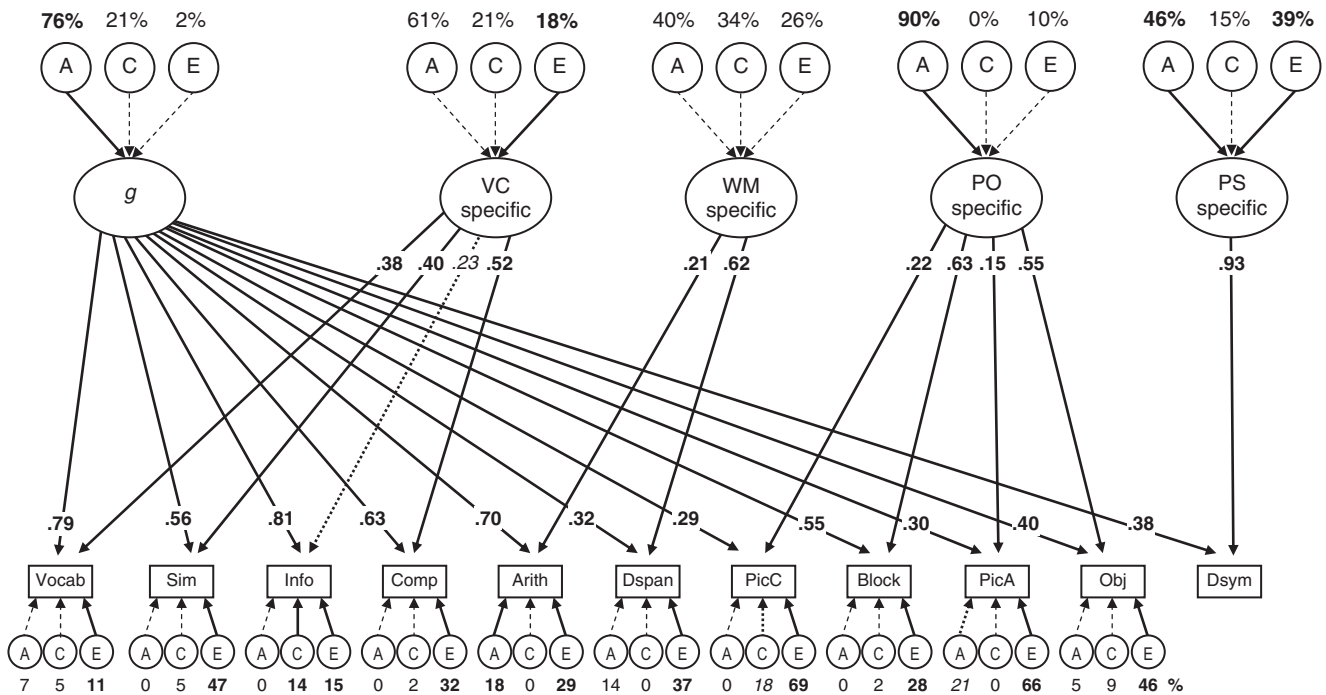
Figure C1. Nested factors model of the subtests of the Wechsler Adult Intelligence Scale—Third Edition. The model depicts a *g* factor, on which all 11 subtests (names abbreviated) load, and four nested factors: Verbal Comprehension (VC) specific, on which Vocabulary, Similarities, Information, and Comprehension load; Working Memory (WM) specific, on which Arithmetic and Digit Span load; Perceptual Organization (PO) specific, on which Picture Completion, Block Design, Picture Arrangement, and Object Assembly load; and Perceptual Speed (PS) specific, which has Digit Symbol Substitution as a single indicator. Because a two-indicator latent variable is not statistically identified when it does not correlate with any other variables in the model (Bollen, 1989), the loadings for the WM-specific variable were fixed at the values estimated in the full model that incorporated correlations with the executive function latent variables. Numbers above the ACEs for the latent variables are the percentages of those variables' variances accounted for by genetic and environmental influences. Numbers occluding arrows are standardized factor loadings. Numbers under the lower ACEs are estimates for task-specific variances. Boldface type and solid lines indicate *p* < .05. Italic type indicates that the .23 loading of the information subtest on the VC-specific variable approached significance (*p* = .066), as did the 18% specific C variance for the picture completion subtest (*p* = .055) and the 21% specific A variance for the picture arrangement subtest (*p* = .061).