Developmental Differentiation of Executive Functions on the NIH Toolbox Cognition Battery

Natacha Akshoomoff[1,2], **Timothy T. Brown**[2,3], **Roger Bakeman**[4], and **Donald J. Hagler Jr.**[5]
Pediatric Imaging, Neurocognition, and Genetics Study

[1]Department of Psychiatry, University of California, San Diego, La Jolla, CA

[2]Center for Human Development, University of California, San Diego, La Jolla, CA

[3]Department of Neurosciences, University of California, San Diego, La Jolla, CA

[4]Department of Psychology, Georgia State University, Atlanta, GA

[5]Department of Radiology, University of California, San Diego, La Jolla, CA

Abstract

Objective: The NIH Toolbox Cognition Battery (NTCB) is a brief computerized method for evaluating neuropsychological functions in children, adolescents, and adults. We examined how performance on the two executive function measures of cognitive flexibility and inhibitory control was related to performance on the other NTCB measures across development.

Method: Participants were 1020 typically developing individuals between the ages of 3 and 21 from the Pediatric Imaging, Neurocognition, and Genetics Study who were divided into five age groups (3–6, 7–9, 10–13, 14–17, and 18–21). Scores were adjusted for sex, level of parental education, and family income.

Results: Although the correlations between the two executive function measures were moderate and consistent across age groups, their correlations with the other five cognitive measures were highest in the youngest age group and decreased across the older age groups. Exploratory factor analysis revealed that all NTCB measures loaded onto a single factor for the 3 to 6 year-olds. Across the older age groups the executive function and processing speed measures loaded onto one factor and the vocabulary knowledge, oral reading, and working memory measures loaded onto a second factor.

Conclusions: These results indicate that younger children's performance on the NTCB is more intercorrelated and less differentiated, while performance on the NTCB executive function

Corresponding Author: Natacha Akshoomoff, Ph.D., UC San Diego Department of Psychiatry, School of Medicine and Center for Human Development, 9500 Gilman Drive, La Jolla, CA 92093-0115, Telephone: 858-822-2757, FAX: 858-822-1602, nakshoomoff@ucsd.edu.

measures becomes more differentiated from performance on the other measures with development. These results support the hypothesis that executive functions become increasingly differentiated from other cognitive functions with development as the functional specialization of neural systems progresses throughout childhood and young adulthood.

Keywords

Computerized Assessment; Cognitive Development; Executive Functions

The NIH Toolbox Cognition Battery (NTCB) was designed to tap key functions (executive function, attention, episodic memory, working memory, language, and processing speed) across the lifespan (ages 3 to 85 years; Gershon et al., 2010). For pediatric studies, this approach has the advantage of a brief, computerized assessment using the same set of measures with young children, older children, and adolescents.

In a previous report, we described the age-related changes in performance on the NTCB from a large normative sample of 1020 individuals ranging in age from 3 to 21 years (Akshoomoff et al., 2014). These data were collected as part of the Pediatric Imaging, Neurocognition, and Genetics (PING) Study (Jernigan et al., 2016). As expected, age accounted for a large portion of the variability in scores. Overall results were very similar to those reported for a smaller sample of children and adolescents in the validation study (Weintraub et al., 2013).

Performance on some executive function measures improves rapidly with age in young children while performance on "higher order" or more complex tasks does not peak until adolescence or early adulthood, particularly those requiring impulse control or the "cognitive control system" (Casey, Giedd, & Thomas, 2000). However, few published behavioral studies to date have characterized the developmental time course of the same set of executive function tasks across the full age range over which they are believed to mature, namely from preschool ages into young adulthood (Best & Miller, 2010). Certain elements of task performance may reflect a broader array of fundamental skills in younger children. Executive functions are characterized by both unity and diversity (Teuber, 1972). Adult studies have shown both a task-domain general performance factor that cannot be explained by general cognitive abilities (unity) as well as task-domain specific factors (diversity; Miyake & Friedman, 2012; Miyake et al., 2000). Studies of executive functions in younger children support a single unitary factor (Nelson et al., 2016; Wiebe, Espy, & Charak, 2008; Wiebe et al., 2011) with emergence of diversity beginning in middle childhood (Huizinga, Dolan, & van der Molen, 2006).

Executive functions may become increasingly differentiated from other cognitive functions with development as the functional specialization of neural systems progresses throughout childhood and adolescence (Johnson, 2011). This differentiation may also reflect the refinement of specific cognitive skills through experience and opportunity (Zelazo et al., 2013). In the validation study conducted by the NTCB developers, they reported that the correlation between both the executive function measures and the receptive vocabulary measure (which was used as a proxy for general intellectual level) declined with age (Zelazo et al., 2013). They suggested that these results reflect not only increasing differentiation

between executive functions and receptive vocabulary ability with age, but also that specific domains of cognitive functioning are less defined in younger children. Further, executive functions in younger children should be less differentiated from other cognitive abilities because of the substantial development of frontal lobe structure and function that occurs throughout childhood, adolescence, and into early adulthood (Mungas et al., 2013). In the NTCB validation study, results from the factor structure (Mungas et al., 2013) and the composite scores (Akshoomoff et al., 2013) provided additional support for a new hypothesis that neurocognitive development involves greater functional specialization of both neural systems and cognitive functions.

Here we examined associations among NTCB variables within five age groups in the PING sample. We defined the age groups to reflect commonly understood developmental periods (early childhood, middle childhood, puberty, adolescence, and young adulthood) by grouping participants ages 3 through 6, 7 through 9, 10 through 13, 14 through 17, and 18 through 21 together. We were particularly interested in how performance on the two primary measures of executive functioning, the Dimensional Change Card Sort Test (a measure of cognitive flexibility) and the Flanker Inhibitory Control and Attention Test (a measure of inhibitory control in the context of selective visual attention), was related to performance on the other five cognitive measures. In the original validation study, the correlations between these executive function measures and the NTCB receptive vocabulary measure (Picture Vocabulary Test) were lower in the 8–15 year olds compared with the 3–6 year-olds (Zelazo et al., 2013). We predicted that, while cognitive flexibility (the Dimensional Change Card Sort) would correlate with inhibitory control (the Flanker Inhibitory Control and Attention Test) in all age groups examined, each would correlate with other tests in younger groups but correlate increasingly less in older age groups, indicating greater differentiation. We also predicted that exploratory factor analyses would show a single factor for younger age groups, indicating that younger children's performance across the NTCB measures was relatively undifferentiated, but that similar factor analyses for older age groups would show two or more factors, indicating greater differentiation of cognitive abilities with age.

Method

Participants

The 1020 participants in this study (486 females and 534 males aged 3 to 21 years) are the same as those recruited for an earlier study reported in (Akshoomoff et al., 2014). All participants were recruited through local postings and outreach activities conducted in the greater metropolitan areas of Baltimore, Boston, Honolulu, Los Angeles, New Haven, New York, Sacramento, and San Diego. The human research protections programs and institutional review boards at the 9 institutions participating in the PING project approved all experimental and consenting procedures. For individuals under 18 years of age, parental informed consent and child assent (for those 7 to 17 years of age) were obtained.

Participants were excluded if there was a reported history of major developmental, psychiatric, or neurological disorders, brain injury, prematurity (i.e., born at less than 36 weeks gestational age), exposure to illicit drugs or alcohol prenatally for more than one trimester, history of head trauma with loss of consciousness for more than 30 minutes, or

other medical conditions that could affect development. Individuals with contraindications for MRI studies (such as dental braces, metallic or electronic implants, claustrophobia, or pregnancy) were also excluded from participating. Individuals with identified or suspected learning disability or ADHD were not excluded since these syndromes are fairly common in pediatric populations.

Information about socioeconomic status (SES) for each participant was based on the parent's indication of 'highest level of parental education' (highest level among those reported for either parent or guardian) and 'family annual income'. Information about race and ethnicity was also collected on the PING Study Demographics and Child Health History Questionnaire. Among the participants who endorsed a single racial category, 56% were White, 13% were Black, and 8% were Asian. The remaining 24% indicated more than one racial category or "Other". Across this sample, 24% of the participants indicated that they were Hispanic/Latino.

For subsequent analyses, participants were divided into five groups by age (3–6, 7–9, 10–13, 14–17, and 18–21 years). The 3–6 year group included thirteen 3-year olds and the 18–21 age group included four 21-year olds; otherwise, ages were relatively evenly distributed within the age groups. Demographic statistics for these five age groups are presented in Table 1. Percentage female did not differ significantly across the age groups; $\chi^2(4, N = 1020) = 7.75$, $p = .11$, nor did percentages for annual family income categories; $\chi^2(16, N = 1020) = 19.4$, $p = .25$. But percentages for parental education categories did vary by age group; $\chi^2(12, N = 1020) = 36.9$, $p < .001$. As noted in Table 1, percentages less than expected (expected is the overall percentage) were college graduate for ages 3–6 (18% vs. 28%), advanced degree for ages 10–13 (26% vs. 34%), and some college for ages 18–21 (18% vs. 24%), whereas greater than expected were advanced degree for ages 3–6 (42% vs. 34%), college graduate for ages 10–13 (39% vs. 28%), and high school or less for ages 18–21 (21% vs. 15%), as gauged by adjusted residuals greater than 1.96 absolute (Bakeman & Quera, 2011).

NIH Toolbox Cognition Battery Measures

The validation study version of the NTCB was utilized for this study and comprised seven tests within five major cognitive domains (see Table 2). Details about the development of the test instruments and reliability and validity data for children ages 3 to 15 years are available (Weintraub et al., 2013). For more detailed descriptions of each measure, see Akshoomoff et al. (2014) and Weintraub et al. (2013). These seven tests were:

1. The *Dimensional Change Card Sort Test* (DC) is a measure of cognitive flexibility or set shifting. The card sorting version of this test has been used to study the development of executive functions in childhood (Beck, Schaefer, Pang, & Carlson, 2011; Zelazo, 2006). Two pictures were presented on the touchscreen monitor that varied along two dimensions (shape, color) and participants were asked to quickly match a series of test pictures to the target pictures switching between matching dimensions.

2. The *Flanker Inhibitory Control and Attention Test* (FL) requires participants to focus on a given middle stimulus in a series and respond quickly regarding its

left-right orientation while inhibiting attention to similar or incongruent stimuli flanking it.

3.   For the *Pattern Comparison Test* (PC), participants must quickly decide whether pairs of side- by-side pictures and designs are the same or not.

4.   For the *Picture Vocabulary Test* (VO), participants were presented with an audio recording of a word and four color photos on the computer screen, and told to select the picture that best corresponds to the meaning of the word.

5.   For the *Oral Reading Recognition Test* (RD), participants were asked to read and pronounce letters and words as accurately as possible.

6.   For the *List Sorting Test* (LS), pictures of different foods and animals were presented along with audio recordings of the name of the object; participants were asked to say the items back in size order from smallest to largest, first within a single dimension (i.e., food or animals) and then on two dimensions (i.e., food then animals).

7.   The *Picture Sequence Memory Test* (PSM) involves recalling the order of increasingly longer series of pictured objects and activities presented on the computer screen with corresponding audio-recorded phrases being played; participants were asked to reproduce the sequence of the pictures over two learning trials by touching each of the pictures on the touchscreen and placing them in the correct order.

Based on the method used by the NTCB developers (Mungas et al., 2013; Zelazo et al., 2013), the NTCB variables were recoded prior to analysis using the Blom rank order normalization algorithm in SAS Proc Rank. This resulted in variables with relatively normal distributions and also established a common scale of measurement of all variables.

Data Analysis

Given the significant education effects for age groups reported earlier, and the near-significant effects for sex, scores for the seven normed NTCB variables described in the previous section were regressed on participant sex, level of parent education, and level of family income. Predicted scores were computed for each variable, for each participant, based on the regression model, and the predicted scores were subtracted from the observed scores. These adjusted or residual scores, which control for differences in sex, level of parent education, and level of family income across the age groups, were used for subsequent analyses (analyses using unadjusted scores yielded essentially similar results; level of education and income correlated .58).

Analyses included Pearson product-moment correlations between test scores, *t*-ratios to assess the magnitude of the difference between two correlations with one variable in common in the same sample (McNemar, 1969;p. 158), and exploratory factor analyses (principal-axis factoring, varimax rotation), separately by age group (factors were defined as those with eigenvalues > 1 confirmed with a scree test and included variables with loadings greater than .40; see e.g., Costello & Osborne, 2005).

Results

The magnitude of the correlations between cognitive flexibility (DC) and the other six tests is shown in Figure 1, left side, while the magnitude of the correlations between inhibitory control (FL) and the other six tests is shown in Figure 1, right side. In each figure the top line represents the correlation between cognitive flexibility and inhibitory control. This correlation ranged from .48 to .59 across the age groups (differences between age groups were not significant per Fisher r to z test). In contrast, correlations between cognitive flexibility and the other five tests—processing speed (PC), vocabulary knowledge (VO), oral reading skill (RD), working memory (LS), and episodic memory (PSM)—and between Inhibitory control and these other five tests were highest in the youngest age group and mainly declined in older age groups.

Of major interest was whether the correlation between cognitive flexibility and inhibitory control (DC-FL) differed significantly from the correlations between cognitive flexibility and the other five tests and, likewise, whether this same correlation differed significantly from the correlations between inhibitory control and the other five tests. Our expectation was that these differences would be minimal in the younger age groups, but become increasingly greater in older age groups. This expectation was largely met.

Figure 2 plots t-ratios across the age groups. These t-ratios (essentially z scores in a sample of this size) assess the magnitude of the difference between two correlations with one variable in common in the same sample (McNemar, 1969;p. 158). Figure 2, left side, compares the correlation between cognitive flexibility and inhibitory control with correlations between cognitive flexibility and the other five tests, while Figure 2, right side, compares the same correlation (which can be expressed as either DC-FL or FL-DC) with correlations between inhibitory control and the other five tests. The horizontal line in these figures represents the $p < .01$ value for the t test.

Figure 2 shows that generally the magnitude of the difference between the cognitive flexibility–inhibitory control correlation and the correlations of either of these tests with other tests increased across all age groups, with some exceptions: (a) correlations of both cognitive flexibility and inhibitory control with processing speed (DC-PC, FL-PC) and with episodic memory (DC-PSM, FL-PSM) decreased from 14–17 to 18–21 years of age, (b) correlations of inhibitory control with vocabulary knowledge (FL-VO), oral reading skill (FL-RD), and working memory (FL-LS) likewise decreased from 14–17 to 18–21 years of age, (c) correlations of cognitive flexibility with vocabulary knowledge (DC-VO) and of inhibitory control with processing speed (FL-PC) and with episodic memory (FL-PSM) failed to increase from 3–6 to 7–9 years of age (see Figure 2).

Results of exploratory factor analyses are shown in Table 3. Only one factor was identified with an eigenvalue > 1 for the 3–6 age group. This factor accounted for 55% of the variance (unrotated matrix; rotation not possible with only one factor); unrotated loadings ranged from .59 to .83. Two factors with eigenvalues > 1 were identified for the other age groups, accounting for 56%, 56%, 53%, and 57% of the variance for the 7–9, 10–13, 14–17, and 18–21 age groups, respectively. We defined a factor as including those tests with rotated factor

loadings of .40 or greater. For the last four age groups, one factor consisted of cognitive flexibility (DC), inhibitory control (FL), and processing speed (PC), with loadings that ranged from .41 to .81 (Factor 2 in Table 3). The other factor consisted of vocabulary knowledge (VO), oral reading skill (RD), and working memory (LS) with loadings that ranged from .40 to .84 (Factor 1 in Table 3). The episodic memory measure (PSM) loaded onto Factor 1 only in the 7–9 age group.

Discussion

As predicted, we found that the pattern of correlations between the NTCB measures differed with age from early childhood to young adulthood and was consistent with the hypothesis of increasing differentiation of cognitive functions. Although the correlations between the two executive function measures of cognitive flexibility and inhibitory control (Dimensional Change Card Sort and Flanker) were moderate and consistent across the age groups, their correlations with the other five cognitive measures were highest in the youngest age group (ages 3 to 6) and increasingly less correlated with these other five cognitive measures in the older age groups. Statistical comparison of the bivariate correlations showed that the magnitude of the difference between the correlation of the two executive function measures and the correlation of either of these measures with the other measures generally increased across the age groups.

The results from the exploratory factor analyses also supported our prediction. All of the NTCB measures loaded onto one factor for the 3 to 6 year-olds and this factor accounted for 55% of the variance. Across the four older age groups, the measures of vocabulary knowledge, oral reading skills, and working memory loaded onto one factor while the two executive function measures and the measure of processing speed (Pattern Comparison) loaded onto a second factor. These results also suggest that in typically developing children, these aspects of executive functions (cognitive flexibility and inhibitory control in the context of visual selective attention) become more clearly differentiated from other critical cognitive skills with age. It is not clear why the episodic memory measure (Picture Sequence Memory) did not load significantly onto either factor for the three oldest age groups, in contrast to the working memory measure (List Sorting). Although we have opted not to name the factors in our results, perhaps one factor reflects more strongly verbal skills while the other reflects more fluid/online processing. It may be that older children and young adults rely on both types of skills when performing the Picture Sequence Memory task.

Using the data from the validation study, the NTCB developers conducted a confirmatory factor analysis to evaluate the dimensional structure underlying the NTCB (Mungas et al., 2013). These analyses included the other test measures used to evaluate the convergent and discriminant validity of the NTCB. The purpose of that particular study and the methods used are therefore quite different from the present study. It is interesting that they also found less differentiation in their 3 to 6 year-olds, and the executive function and processing speed measures loaded together on a separate factor in their group of 8 to 15 year-olds.

Although it may be difficult to disentangle certain task demands, such as the executive functioning requirements in some tests of processing speed (Cepeda, Blackwell, &

Munakata, 2013), this is less likely to explain the pattern of results for younger children across all of the NTCB measures. It is likely that factors such as attention, sustained effort, and language have a stronger, more general influence on task performance among younger children (Akshoomoff, 2002).

Our results, though not entirely surprising, have implications for theories of developing executive functions that may relate to changes in the functional brain organization. We found that the youngest age group showed only one factor that accounted for a relatively large proportion of the variance in performance across many different kinds of tasks. This suggests the involvement of overlapping, "general purpose" cognitive systems across different kinds of tasks early in development. This could reflect the fact that many purportedly "non-executive" tasks are nevertheless novel and challenging at younger ages and engage effortful control systems (e.g., prefrontal cortical regions) similar to those required for more classic adult executive tasks until they become more routinized. This would be consistent with some evidence from functional neuroimaging studies of both child development and of adult skill acquisition (Brown et al., 2005; Luna et al., 2001; Raichle et al., 1994).

This interpretation would also be consistent with some aspects of both a skill learning theoretical framework for the developing functional organization, as well as an interactive specialization perspective (Johnson, 2011; Klingberg, 2014). However, functional neuroimaging is required to address the localization aspects of these theories, and available evidence suggests that the developing cognitive specialization takes many forms, including regions involved early that are "tuned", newly-involved regions, especially within frontal cortex, and regions that participate in the same tasks only at younger ages (Brown et al., 2005; Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002; Casey et al., 1995; Schlaggar et al., 2002).

Our behavioral study, nevertheless, addresses an important question about the cognitive structure of developing executive and cognitive control functions. Our results suggest that the relationship between specific task demands and the construct of executive functions changes with development and that with age there is increasing differentiation between traditionally "executive" tasks and other cognitive functions. Although there may be greater shared variance in performance across the NTCB measures in younger children, we are not suggesting that the same cognitive processes are involved across all of these measures.

This study has some limitations. The PING study sample was limited to 9 U.S. locations and is not a nationally representative sample. Overall the sample had a higher level of parental education and annual family income than the general U.S. population of children. These are factors that are known to be associated with many aspects of well-being in children, including test performance. The results of our analyses were the same when adjusted for these demographic factors but future studies of developmental changes in NIH Cognition Toolbox should target a broader sample to examine these factors, as well as race/ethnicity. Not accounting for demographic effects can underestimate or overestimate deviations from expected performance (Casaletto et al., 2015). The PING study did not exclude potential participants with a known or suspected learning disability or ADHD diagnosis. However, no

testing was conducted to screen for ADHD and therefore verification of a diagnosis or identification of additional participants who may have met criteria for a learning disability and/or ADHD was not possible. Additional studies are needed to determine how children with these diagnoses perform on the NIH Cognition Toolbox, particularly the executive function measures.

Acknowledgments

References

Akshoomoff N, Beaumont JL, Bauer PJ, Dikmen SS, Gershon RC, Mungas D, Slotkin J, Tulsky D, Weintraub S, Zelazo PD, & Heaton RK (2013). VII. NIH toolbox cognition battery (cb): composite scores of crystallized, fluid, and overall cognition. Monogr Soc Res Child Dev, 78(4), 119–132. [PubMed: 23952206]

Akshoomoff N, Newman E, Thompson WK, McCabe C, Bloss CS, Chang L, Amaral DG, Casey BJ, Ernst TM, Frazier JA, Gruen JR, Kaufmann WE, Kenet T, Kennedy DN, Libiger O, Mostofsky S, Murray SS, Sowell ER, Schork N, Dale AM, & Jernigan TL (2014). The NIH Toolbox Cognition Battery: Results from a large normative developmental sample (PING). Neuropsychology, 28(1), 1–10. [PubMed: 24219608]

Akshoomoff NA (2002). Selective attention and active engagement in young children. Developmental Neuropsychology, 22, 625–642. [PubMed: 12661973]

Bakeman R, & Quera V (2011). Sequential analysis and observational methods for the behavioral sciences. Cambridge, UK: Cambridge University Press.

Beck DM, Schaefer C, Pang K, & Carlson SM (2011). Executive Function in Preschool Children: Test-Retest Reliability. J Cogn Dev, 12(2), 169–193. [PubMed: 21643523]

Best JR, & Miller PH (2010). A developmental perspective on executive function. Child Development, 81(6), 1641–1660. [PubMed: 21077853]

Brown TT, Lugar HM, Coalson RS, Miezin FM, Petersen SE, & Schlaggar BL (2005). Developmental changes in human cerebral functional organization for word generation. Cerebral Cortex, 15(3), 275–290. [PubMed: 15297366]

Bunge SA, Hazeltine E, Scanlon MD, Rosen AC, & Gabrieli JD (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. Neuroimage, 17(3), 1562–1571. [PubMed: 12414294]

Casaletto KB, Umlauf A, Beaumont J, Gershon R, Slotkin J, Akshoomoff N, & Heaton RK (2015). Demographically Corrected Normative Standards for the English Version of the NIH Toolbox Cognition Battery. J Int Neuropsychol Soc, 21(5), 378–391. [PubMed: 26030001]

Casey BJ, Cohen JD, Jezzard P, Turner R, Noll DC, Trainor RJ, Giedd J, Kaysen D, Hertz-Pannier L, & Rapoport JL (1995). Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. Neuroimage, 2(3), 221–229. [PubMed: 9343606]

Casey BJ, Giedd JN, & Thomas KM (2000). Structural and functional brain development and its relation to cognitive development. Biol Psychol, 54(1–3), 241–257. [PubMed: 11035225]

Cepeda NJ, Blackwell KA, & Munakata Y (2013). Speed isn't everything: complex processing speed measures mask individual differences and developmental changes in executive control. Dev Sci, 16(2), 269–286. [PubMed: 23432836]

Costello AB, & Osborne J (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment Research and Evaluation, 10, 1–9.

Gershon RC, Cella D, Fox NA, Havlik RJ, Hendrie HC, & Wagster MV (2010). Assessment of neurological and behavioural function: the NIH Toolbox. Lancet Neurol, 9(2), 138–139. [PubMed: 20129161]

Huizinga M, Dolan CV, & van der Molen MW (2006). Age-related change in executive function: developmental trends and a latent variable analysis. Neuropsychologia, 44(11), 2017–2036. [PubMed: 16527316]

Jernigan TL, Brown TT, Hagler DJ, Jr., Akshoomoff N, Bartsch H, Newman E, Thompson WK, Bloss CS, Murray SS, Schork N, Kennedy DN, Kuperman JM, McCabe C, Chung Y, Libiger O, Maddox M, Casey BJ, Chang L, Ernst TM, Frazier JA, Gruen JR, Sowell ER, Kenet T, Kaufmann WE, Mostofsky S, Amaral DG, & Dale AM (2016). The Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository. Neuroimage, 124(Pt B), 1149–1154. [PubMed: 25937488]

Johnson MH (2011). Interactive specialization: a domain-general framework for human functional brain development? Dev Cogn Neurosci, 1(1), 7–21. [PubMed: 22436416]

Klingberg T (2014). Childhood cognitive development as a skill. Trends Cogn Sci, 18(11), 573–579. [PubMed: 25042686]

Luna B, Thulborn KR, Munoz DP, Merriam EP, Garver KE, Minshew NJ, Keshavan MS, Genovese CR, Eddy WF, & Sweeney JA (2001). Maturation of widely distributed brain function subserves cognitive development. Neuroimage, 13(5), 786–793. [PubMed: 11304075]

McNemar Q (1969). Psychological statistics. New York: Wiley.

Miyake A, & Friedman NP (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. Curr Dir Psychol Sci, 21(1), 8–14. [PubMed: 22773897]

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. Cogn Psychol, 41(1), 49–100. [PubMed: 10945922]

Mungas D, Widaman K, Zelazo PD, Tulsky D, Heaton RK, Slotkin J, Blitz DL, & Gershon RC (2013). VII. NIH toolbox cognition battery (cb): factor structure for 3 to 15 year olds. Monogr Soc Res Child Dev, 78(4), 103–118. [PubMed: 23952205]

Nelson JM, James TD, Choi HJ, Clark CA, Wiebe SA, & Espy KA (2016). III. Distinguishing executive control from overlapping foundational cognitive abilities during the preschool period. Monogr Soc Res Child Dev, 81(4), 47–68. [PubMed: 27943324]

Raichle ME, Fiez JA, Videen TO, MacLeod AM, Pardo JV, Fox PT, & Petersen SE (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. Cerebral Cortex, 4(1), 8–26. [PubMed: 8180494]

Schlaggar BL, Brown TT, Lugar HM, Visscher KM, Miezin FM, & Petersen SE (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. Science, 296(5572), 1476–1479. [PubMed: 12029136]

Teuber H-L (1972). Unity and diversity of frontal lobe functions. Acta Neurobiologiae Experimentalis, 32, 615–656. [PubMed: 4627626]

Weintraub S, Bauer PJ, Zelazo PD, Wallner-Allen K, Dikmen SS, Heaton RK, Tulsky DS, Slotkin J, Blitz DL, Carlozzi NE, Havlik RJ, Beaumont JL, Mungas D, Manly JJ, Borosh BG, Nowinski CJ, & Gershon RC (2013). I. Nih toolbox cognition battery (cb): introduction and pediatric data. Monogr Soc Res Child Dev, 78(4), 1–15.

Wiebe S, Espy K, & Charak D (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. Developmental Psychology, 44, 575–587. [PubMed: 18331145]

Wiebe SA, Sheffield T, Nelson JM, Clark CA, Chevalier N, & Espy KA (2011). The structure of executive function in 3-year-olds. [Research Support, N.I.H., Extramural]. Journal of Experimental Child Psychology, 108(3), 436–452. [PubMed: 20884004]

Zelazo PD (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. Nat Protoc, 1(1), 297–301. [PubMed: 17406248]

Zelazo PD, Anderson JE, Richler J, Wallner-Allen K, Beaumont JL, & Weintraub S (2013). II. NIH toolbox cognition battery (cb): measuring executive function and attention. Monogr Soc Res Child Dev, 78(4), 16–33. [PubMed: 23952200]

Public Significance Statement:

The NIH Toolbox Cognition Battery is a brief computerized method for evaluating neuropsychological functions in children, adolescents, and adults using the same set of tests. Characterization of performance on these measures across development will help us understand typical cognitive and brain development in children and apply these results to studies of school performance and children with neurodevelopmental disorders.
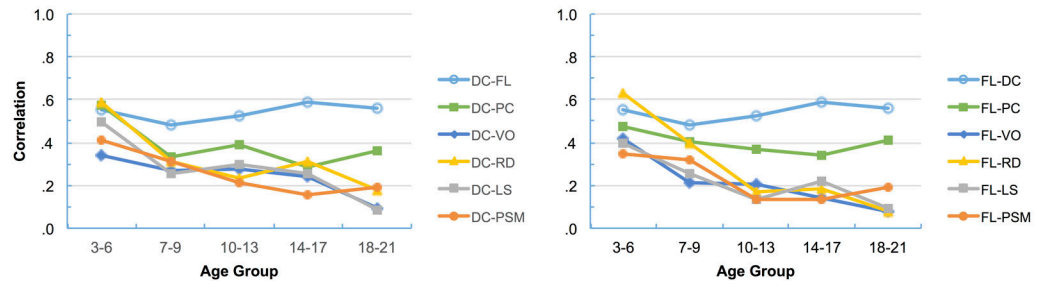
Figure 1.
Correlations of cognitive flexibility (DC, left side) and inhibitory control (FL, right side) with each other and with processing speed (PC), vocabulary knowledge (VO), oral reading skill (RD), working memory (LS), and episodic memory (PSM).
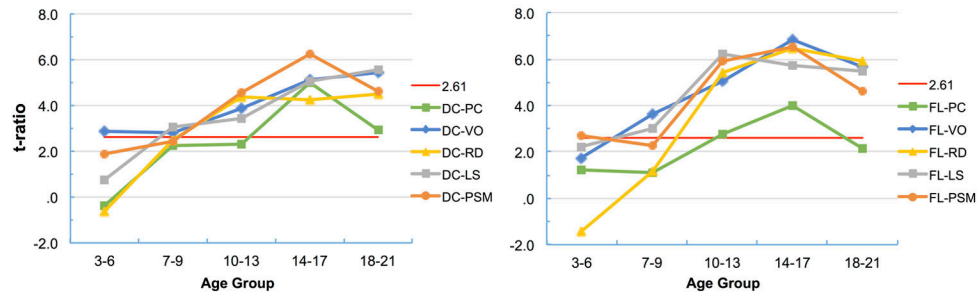
Figure 2.
t-ratios gauging the magnitude of the difference between the cognitive flexibility–inhibitory control correlation and correlations of cognitive flexibility (DC, left side) and inhibitory control (FL, right side) with processing speed (PC), vocabulary knowledge (VO), oral reading skill (RD), working memory (LS), and episodic memory (PSM).

Table 1.

Demographic Characteristics of the Study Sample

		Age group (years)				
Variable	All	3–6	7–9	10–13	14–17	18–21
N	1020	148	207	262	229	174
Mean age (years)	12.5	5.4	8.5	12.0	15.8	19.5
Female (%)	48	50	50	42	45	55
Parental education						
High school or less	15	14	14	12	14	21$^+$
Some College	24	27	25	24	26	18$^-$
College Graduate	28	18$^-$	24	39$^+$	28	24
Advanced degree	34	42$^+$	36	26$^-$	33	37
Annual family Income						
Less than 10K	13	12	13	10	11	19
10K–49K	18	21	20	17	16	20
50K–99K	30	34	28	32	32	23
100K–149K	19	20	18	21	19	17
150K or more	20	13	21	21	22	21

Note. Percentages for parental education and annual family income may not sum to 100 due to rounding. Percentage female and percentages for family income categories did not differ significantly across the age groups, but percentages for education categories did. For education, − and + superscripts indicate age group percentages significantly below and above expected (i.e., overall percentages), as gauged by adjusted residuals less than −1.96 and greater than +1.96, respectively; see text for details.

Table 2.

NIH Toolbox Cognition Battery Measures

Domain and ability	Toolbox mnemonic and test
Executive Function: Cognitive Flexibility	DC: Dimensional Change Card Sort
Executive Function: Inhibitory Control	FL: Flanker Inhibitory Control and Attention
Processing Speed	PC: Pattern Comparison
Language: Vocabulary Knowledge	VO: Picture Vocabulary
Language: Oral Reading Skill	RD: Oral Reading Recognition
Working Memory	LS: List Sorting
Episodic Memory	PSM: Picture Sequence Memory

Note. Ability is given if different from domain.

Table 3.

Factor Analysis Loadings for Each Age Group

Test	Ability	Age group (years)								
		3–6	7–9		10–13		14–17		18–21	
		Factor 1	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
DC	Cognitive Flexibility	**.73**	.34	**.55**	.23	**.74**	.28	**.68**	.10	**.69**
FL	Inhibitory Control	**.70**	.34	**.65**	.11	**.68**	.12	**.81**	.03	**.79**
PC	Processing Speed	**.64**	.00	**.62**	.20	**.49**	.08	**.41**	.10	**.52**
VO	Vocabulary Knowledge	**.59**	**.55**	.07	**.81**	.12	**.77**	.04	**.73**	.04
RD	Oral Reading Skill	**.83**	**.78**	.15	**.63**	.14	**.66**	.13	**.84**	.08
LS	Working Memory	**.66**	**.48**	.18	**.49**	.20	**.40**	.25	**.51**	.14
PSM	Episodic Memory	**.64**	**.41**	.30	.34	.17	.23	.17	.24	.26

Note. Scores are unrotated factor loading for the 3–6 year age group, rotated factor loading otherwise, from exploratory factor analyses. Loadings of .40 or greater are bolded. Only one factor was extracted for the 3–6 year age group, two factors for all other age groups.