Unity and Diversity of Executive Function in Middle Childhood: Based on Latent Variable Analysis and Network Analysis

Authors: Huang Bijuan, Zhu Hongxiang, Liu Chang, Luo Ruifan, Hongmin Feng, Li Hongxia, Si Jiwei, Si Jiwei

Date: 2023-11-03T09:58:53+00:00

Abstract

As a higher-order cognitive function in humans, executive function constitutes a core element that facilitates individual academic success and physical and mental well-being. However, current research on executive function measurement has emphasized single tasks, lacking an integrated perspective on its structure and stability. This study employs a longitudinal design, conducting a one-year follow-up investigation of 756 children in grades three to four (mean age = 9.25, 51.85% female), and investigates the structure and stability of executive function in middle childhood through latent variable analysis and network analysis. Results indicate that executive function in middle childhood manifests as a common factor of inhibitory control and cognitive flexibility/working memory (i.e., the "I + W/C" model); six executive function tasks were identified as constituting a stable component; and the unity (general factor) and diversity (component-specific factors) of executive function during this period were relatively stable. These findings underscore the necessity of simultaneously considering both the unity and diversity of executive function, advancing a profound understanding of the developmental structure of executive function in children.

Full Text

Unity and Diversity of Children's Executive Functions During Middle Childhood: Latent-Variable Analysis and Network Analysis

Bijuan Huang[1,2], Hongxiang Zhu[1], Chang Liu[1], Ruifan Luo[1], Hongmin Feng[1], Hongxia Li[1], Jiwei Si[1,*]

[1]Research Center for Child Development and Education, Shandong Normal University, Jinan, 250014, China

[2]Institute of Brain Science and Brain-Inspired Research, Shandong First Medical University (Shandong Academy of Medical Sciences), Jinan, 2550117, China

This research was supported by the National Natural Science Foundation of China General Program "The Gene-Brain-Behavior Study on Individual-Environment Interaction in the Development of Children's Math Anxiety" (Project No.: 31971010).

Abstract

As high-level cognitive processes, executive functions (EF) refer to a set of top-down neurocognitive processes that serve conscious, goal-directed control of thought, action, and emotion, which are crucial for children's academic success and mental and physical health. Executive functions include inhibitory control, working memory, and cognitive flexibility. Although abundant theories and studies exist, several limitations remain. First, previous studies have measured EF using single tasks, which are rarely capable of decomposing different components of EF. They lack systematic investigation of the unity and diversity of EF from an integrative perspective. Second, few studies have used longitudinal designs to examine the stability of the unity and diversity of children's executive functions during middle childhood. The present study attempted to address these issues.

From the perspective of a unity/diversity framework of EF, combined with latent-variable analysis and network analysis, the present study followed 756 students from grades 3 to 4 (Mage = 9.25 years, 51.85% girls) to systematically and comprehensively explore the unity and diversity of EF during middle childhood. Participants were recruited to complete six tasks related to the three main components of EF.

The results showed that: The structure of children's EF in middle childhood included inhibitory control and a combined factor of working memory and cognitive flexibility (i.e., the "I + W/C" model); these six tasks were organized into one stable component; and the unity (general factor) and diversity (component-specific factors) of EF remained relatively stable during this period. These findings highlight the importance of understanding both the unity and diversity of EF among middle childhood from a developmental perspective and provide new insights into the measurement of EF.

Keywords: executive functions, unity, diversity, middle childhood

1.1 Assessment of Executive Functions

Currently, there are two common approaches to assessing executive functions: performance-based tests (e.g., inhibitory control assessed via Go/Nogo tasks) and standardized questionnaire assessments (e.g., BRIEF) (Isquith et al., 2013; McAuley et al., 2010; Strauss et al., 2006; Toplak et al., 2013). Below, we briefly introduce these two main assessment approaches.

Standardized questionnaire assessment tools mainly include BRIEF, BDEFS-CA, D-REF, BASC-2, and CEFI (Goldstein & Naglieri, 2014). Among them, BRIEF is currently the most commonly used standardized questionnaire assessment. BRIEF-2 is suitable for assessing executive function behavioral performance in children aged 5–18 years, evaluated by parents or teachers based on children's daily behavioral performance, measuring behavioral regulation (inhibition, shifting, and emotional control) and metacognitive problem-solving (initiation, task organization/planning, environmental organization, self-monitoring, and working memory). However, standardized questionnaire assessments are not the optimal form for evaluating children's executive functions for the following reasons: On the one hand, teacher or parent reports of executive functions cannot effectively reflect children's actual status; on the other hand, standardized questionnaire assessments have relatively more items, and children have limited language abilities and poor motivation to respond (Shi et al., 2019). Children's executive function tests need to be simple and easy to understand, relatively independent of language ability, sufficiently simplified to maintain children's motivation, and even interesting.

Compared with standardized questionnaire assessments, performance-based tests are more suitable for evaluating children's executive functions. Performance-based tests focus on general abilities, more precisely record the speed and accuracy of children's responses, and reduce dependence on language instructions and feedback (De Luca et al., 2003; Wild & Musser, 2014). Additionally, they can capture specific components of executive functions. When exploring theoretical issues in EF development, performance-based tests are necessary (Shi et al., 2019). Performance-based tests include test batteries and individual tasks (McCoy, 2019). Commonly used test batteries in research include CANTAB, D-KEFS, TEC, and NIHTB-CB (Goldstein & Naglieri, 2014). Among them, NIHTB-CB is a relatively comprehensive neurobehavioral measurement method that can quickly assess cognitive, emotional, sensory, and motor functions, suitable for participants aged 3–85 years. Individual test tasks aim to focus on a specific component of EF. Working memory test tasks mainly include the backward digital span from the Wechsler Intelligence Test, N-back paradigm, etc.; inhibitory control test tasks mainly include Oddball, Go/Nogo, Stop-Signal paradigms; cognitive flexibility test tasks mainly include WCST (Wisconsin Card Sorting Test), TMT (Trail Making Test), etc.

These two assessment approaches provide different but important information (Toplak et al., 2013). Given that this study examines middle childhood children and focuses on investigating the structure and stability of children's executive functions, performance-based tests were ultimately selected to assess children's executive functions.

1.2 Constructing the Structure of Executive Functions

In some studies that use multiple tasks to measure executive functions, although correlations between tasks are low or non-significant (Huizinga et al., 2006), latent variable analysis can be used to identify the latent structure underlying observed cognitive task performance (Gorsuch, 1983; Karr et al., 2018; Wiebe et al., 2008). For example, exploratory factor analysis can extract multiple latent factor structures (Friedman & Miyake, 2017; Karr et al., 2018; Miyake et al., 2000). When assessing individual executive functions, performance-based tests face a major challenge—task impurity, meaning that most tests involve non-executive function processing (Karr et al., 2018; Miyake & Friedman, 2012; Miyake et al., 2000; Toplak et al., 2013). For example, color naming and articulation speed in the Stroop task (Miyake & Friedman, 2012). Using confirmatory factor analysis (CFA) to extract the common variance of these tasks can minimize task impurity (e.g., Huizinga et al., 2006; Miyake et al., 2000). By extracting relatively "clean" latent variables, the components identified and constructed through confirmatory factor analysis are considered the underlying cognitive structures of different task performances (Miyake et al., 2000). An increasing number of developmental studies have used confirmatory factor analysis to explore the developmental changes in different components of EF in preschool children (e.g., Carlson et al., 2014; Clark et al., 2013; Espy et al., 1999; Garon et al., 2008; Wiebe et al., 2008), school-age children (e.g., Huizinga et al., 2006; Lehto et al., 2003), adolescents (e.g., Feng et al., 2022; Friedman et al., 2011; Xu et al., 2013), and adults (e.g., Feng et al., 2022; Friedman & Miyake, 2004; Miyake et al., 2000, 2001).

Unlike latent variable analysis, network analysis considers how variables are associated after accounting for all commonalities among them. Through data-driven methods, network analysis can provide new insights for identifying unique and common cognitive mechanisms and is commonly used to understand cognitive structures, particularly complex and interrelated components such as intelligence, psychopathology, and personality. Younger et al. (2023) first applied this method to analyze EF structure, finding that stable components could be identified from California children's performance on eight tasks at least until age 10 (fifth grade), and that the separation of these components persisted at least until age 14 (eighth grade). This method examines the structure of EF from a holistic perspective rather than constructing and comparing multiple competing models, better reflecting the organizational characteristics of the data, and can serve as a complement to latent variable analysis (Cai et al., 2019; Younger et al., 2023).

1.3 The Structure of Executive Functions—Unity and Diversity

Executive functions emerge early in life and continue to develop through middle to late adolescence or early adulthood (Huizinga et al., 2006; McAuley et al., 2010). Their structure gradually differentiates from the preschool stage (Best & Miller, 2010), transforming from a single structure to a series of diverse, interactive processing processes that are currently widely accepted (Karr et al., 2018). Garon et al. (2008) proposed an integrative model of executive functions based on the continuity and hierarchical nature of cognitive skill development, suggesting that EF components develop from simpler cognitive skills or result from the coordination of simpler cognitive skills. These simpler cognitive skills are crucial for educational development, such as goal maintenance, and are necessary conditions for effectively performing various tasks; while combination skills are skills related to the actual components of EF or central executive functions. Moreover, these skills become increasingly apparent and mature with individual brain development. This model emphasizes the non-negligible role of EF unity. In both child and adult populations, different EF test tasks are interrelated; more importantly, all latent factors are correlated across different age groups, indicating the existence of a common process (Friedman & Miyake, 2004; Hughes & Devine, 2019; Lehto et al., 2003; Miyake et al., 2000).

The unity/diversity framework of executive functions (Miyake & Friedman, 2012) posits that each executive function (e.g., working memory) can be decomposed into a universal component or unity (corresponding to a general factor) and a unique component or diversity (corresponding to specific factors, such as working memory-specific factors). Through confirmatory factor analysis methods, tasks related to EF components can be more purely mapped onto latent cognitive processes (including EF general and specific factors). Friedman and colleagues argue that the general factor in a bifactor model can represent the unity of EF, and this model can more directly and clearly characterize the relationship between unity and diversity (Friedman & Miyake, 2017; Friedman et al., 2008). However, Karr et al. (2018) summarized 40 studies that used confirmatory factor analysis to explore EF, re-analyzing results from 46 samples (N = 9,756) and found that the best model for EF differs across developmental stages: preschool children more often support single-factor or two-factor models, school-age children more often support three-factor models, and adolescents/adults support three-factor or bifactor models. That is, bifactor models may not exist in preschool and school-age children, and general and specific factors in bifactor models cannot be extracted. To better depict the unity and diversity of EF in middle childhood, when bifactor models cannot achieve good fit with behavioral data, this study suggests that the general factor from a single-factor model can be extracted as an indicator of unity, while other factors directly related to EF components serve as indicators of diversity. To our knowledge, no study has directly examined the structure of EF in middle childhood children in the context of Chinese educational culture, and it remains uncertain whether the structure of EF in Chinese children at this stage is similar to that in Western populations.

1.4 Purpose of This Study

By reviewing domestic and international progress on EF-related topics, the following issues urgently need to be addressed in current research: On the one hand, EF measurement indicators are not pure; domestic studies commonly use single tasks to measure EF components, which are affected by task impurity and cannot extract "clean" components. On the other hand, there is a lack of attention to the structure and stability of EF in middle childhood children in the context of Chinese culture. Although researchers have described and explored the developmental trends of EF components (mainly inhibitory control, working memory, and cognitive flexibility), which have important reference value for understanding cognitive development characteristics at different stages—for example, Qi et al. (2021) summarized the developmental characteristics of different EF components—the structure of EF gradually differentiates from the preschool stage (Best & Miller, 2010). Individual difference studies on EF have also found that EF has unity and diversity (Miyake & Friedman, 2012; Miyake et al., 2000), and its structure differs across age stages (Karr et al., 2018). For instance, Brydges et al. (2014) found that single-factor and two-factor models of EF exist in middle to late childhood. Therefore, the unity and diversity of EF in middle childhood children need to be further explored.

Reviewing current EF development research reveals the following limitations: (1) Different studies use different tasks to assess the same component. For example, when assessing inhibitory control, Go/Nogo or Flanker tasks are used; (2) Using single tasks to measure a component cannot extract effective "clean" components; (3) Most involve cross-sectional designs, with only a limited number of longitudinal studies (e.g., Brydge et al., 2014; Friedman et al., 2011, 2016), which ignore the longitudinal invariance of EF and cannot determine which factor structures are stable and replicable; (4) Most studies have small sample sizes, generally 100–300 (e.g., Brydge et al., 2014; N = 135), which cannot obtain stable effect sizes, and latent variable analysis has many parameters requiring larger sample sizes to obtain stable structures; (5) There is a lack of effective indicators to measure the unity and diversity of EF in middle childhood children; the two are not mutually exclusive and both play important roles in child development, but their contributions cannot be examined simultaneously.

Given the above issues and limitations, this study combines two models of EF unity/diversity, adopts a one-year interval longitudinal design from a behavioral level, assesses a large sample of children's EF through performance-based multitask testing, and identifies the structure and stability of EF in middle childhood children in the context of Chinese culture through latent variable analysis and network analysis. This study selected third grade as the starting point for tracking based on the following considerations: (1) Childhood and adolescence are periods of rapid physical and mental development, and this stage is also an important period for cognitive control and emotional development. As a high-level psychological function, the normal development of cognitive control ensures individuals have sufficient subjective agency and self-control, and plays a mutual constraining and promoting role with other systems (Qi et al., 2021); (2) Different components of EF in children at this stage have further developed (Lehto et al., 2003), and related research in this area will focus on avoiding risk factors and enhancing protective factors for EF, providing more reference value for understanding and training EF.

2.1 Participants

The data for this study were selected from an ongoing two-year longitudinal research project that primarily focuses on whether and how domain-general cognitive abilities play a role in the development of math anxiety in middle childhood children. The project recruited third to fourth grade participants from three public primary schools in Jinan. Data collection was conducted from the fall semester of 2021 to the spring semester of 2023, with each wave spaced six months apart. This study selected data from the first and third waves for analysis. The project was approved by the local university ethics committee. Parental and school informed consent was obtained before each survey.

A total of 1,192 children participated in the first wave (98.07% completed more than three tasks and met the criteria). Due to pandemic prevention and control reasons, only 956 children participated in the third wave (91.53% completed more than three tasks and met the criteria), with 1,383 children participating across both waves. Due to objective factors such as school transfers (including transfers out and new transfers), absences, and dropouts, not all children participated in all time points. Children who did not provide any qualified data on tasks at both time points were excluded from the analysis (N = 510); children who did not effectively complete more than half of all tasks at either time point were excluded (N = 105); children with a Raven's intelligence test grade of five were excluded (N = 12). The final number of valid participants for data analysis was 756, of which 312 provided valid data on all tasks at all time points. Children with complete data did not differ significantly from those with incomplete data in handedness ($^2 = 0.32$, $p > 0.05$), but differed significantly in age ($t = 3.07$, $p = 0.002$), gender ($^2 = 6.48$, $p = 0.011$), and main variables ($|t|$s ≥ 2.07, ps < 0.05), with children with complete data performing better at all time points. Little's MCAR test results were significant ($p < 0.001$), indicating that data missingness was not completely random. Basic information for the 756 children included in the final analysis is shown in Table 1 . The mean age was 9.25 years, with 76.04% of fathers and 70.60% of mothers having education above high school level.

Table 1 Demographic Information of Participants

Variable	N = 756
Age (years), M (SD)	9.25 (0.56)
Age range	8.17–10.75

Variable	N = 756
Gender (male, %)	364 (48.15%)
Ethnicity (Han, %)	750 (99.21%)
Handedness (right-handed, %)	746 (98.68%)
IQ (%)	
5–25	201 (26.59%)
25–75	420 (55.56%)
75–95	124 (16.40%)
95–100	11 (1.45%)
Parental education (father/mother, %)	N = 722/721
Primary school or below	1.39% / 2.08%
Junior high school (including incomplete)	22.58% / 27.32%
High school or technical secondary school (including incomplete)	29.92% / 31.21%
College (including night school, TV university, party school)	27.01% / 21.64%
Graduate school (master's or doctoral)	18.14% / 17.06%
Postgraduate (master's or doctoral)	0.97% / 0.69%

2.2.1 Executive Function Tests

A behavioral test battery was used to assess executive functions. The battery included six tasks: two working memory tasks, two inhibitory control tasks, and two cognitive flexibility tasks. All tasks were run on a computer, with stimulus materials presented on a 15-inch Lenovo monitor (resolution 1,024 × 68, refresh rate 60 Hz), with a gray background and white foreground. All fonts were Times New Roman, size 32.

(1) Inhibitory Control

Following Maasalo et al. (2021), a revised, child-friendly version of the Go/Nogo task and Oddball task (Downes et al., 2017; Seer et al., 2016) was used to assess children's inhibitory control ability, measuring individuals' ability to inhibit dominant responses.

Oddball Task. Different-sized black beans were presented at the center of the screen (large black bean visual angle approximately 1.15° × 2.29°; small black bean visual angle approximately 0.46° × 0.92°). Children were required to press the "F" key with their left index finger as quickly and accurately as possible when they saw small black beans (150 trials, accounting for 75% of total trials), and press the "J" key with their right index finger when they saw large black beans (50 trials, accounting for 25% of total trials). Children's response accuracy and reaction time were recorded. Large black bean trials were target trials. Hit rate (correct key press on target trials), false alarms (incorrect key press on non-target trials), and discriminability (d' = Z hit rate − Z false alarm rate) were calculated. Discriminability d' served as the behavioral indicator of children's inhibitory control ability, with higher scores indicating better discriminability.

Go/Nogo Task. Dots of different colors (visual angle approximately 1.83° × 1.83°) were presented at the center of the screen and were divided into Go stimuli and Nogo stimuli based on color. Children were required to press the "F" key as quickly and accurately as possible when they saw Go stimuli (150 trials, accounting for 75% of total trials), and not to respond when they saw Nogo stimuli (50 trials, accounting for 25% of total trials). Children's response accuracy and reaction time were recorded. Nogo trials were target trials. Hit rate, false alarms, and discriminability d' were calculated. Discriminability d' served as the behavioral indicator of children's inhibitory control ability.

The procedure parameters for the two tasks were identical: a fixation point was presented for 500 ms, followed by stimulus presentation for 500 ms, then a 1000 ms blank screen; the inter-trial interval was 800–1000 ms to reduce children's anticipation (detailed procedure shown in Figure 1 [Figure 1: see original paper]A and Figure 1B). Children were required to respond during the stimulus presentation screen and the subsequent blank screen (total 1500 ms). Before the formal experiment, children completed 10 practice trials to ensure understanding of the experimental procedure, with a correct rate of 90% required to proceed to the formal experiment. Each task included 4 blocks, with each block containing 50 trials and brief rest periods between blocks. Each task lasted approximately 5 minutes.

(2) Working Memory

Following Maasalo et al. (2021), the 1-back task from the N-back task (Pelegrina et al., 2020; Yaple & Arsalidou, 2018) and the Digital Span subtest from the Wechsler Intelligence Test (Digital Span; Wechsler, 2002) were used to assess children's working memory capacity.

1-back Task. A numerical 1-back task was used, with stimuli being 9 Arabic digits (1–9). In each trial, a digit was presented at the center of the screen for 400 ms, followed by a 1000 ms blank screen. Children were required to make a key press response during the stimulus presentation screen and blank screen (total 1400 ms); the inter-trial interval was 500 ms (detailed procedure shown in Figure 1C). Children needed to judge whether the digit they saw was the same as the previous digit, pressing the "F" key if it was the same and the "J" key if it was different. Before the formal experiment, children completed 10 practice trials to ensure understanding of the experimental procedure, with a correct rate of 90% required to proceed to the formal experiment. The task was divided into 4 blocks, with each block containing 50 trials and brief rest periods between blocks. Digit-same trials were target trials, accounting for 25% of total trials. The task lasted 5 minutes. Children's response accuracy and reaction time were recorded. Hit rate, false alarms, and discriminability d' were calculated, with discriminability d' serving as the behavioral indicator of children's working memory ability.

Backward Digital Span Task. A series of digits (digit sets) ranging from 1–9 were presented to children. Digit set lengths ranged from 2–6, with 6 sets of each length. Digits in each set were played at a rate of one per second, with digit sets presented to participants in ascending order of length. After each digit set was played, children were required to repeat the digits they heard in the headphones in reverse order. When children correctly repeated 4 consecutive sets of the same length, they proceeded to the next length; when they incorrectly repeated 3 consecutive sets of the same length, the test was discontinued. The number of correctly repeated sets was recorded as the indicator of digit span, with higher scores indicating better digit working memory ability.

(3) Cognitive Flexibility

Following Maasalo et al. (2021), a computer version of the WCST (Lange et al., 2017) and TMT (Takacs & Kassai, 2019) tasks were used to assess children's cognitive flexibility.

WCST. Sixty-four cards were created based on shape (square, circle, pentagram, and cross), color (red, green, yellow, and blue), and quantity (1, 2, 3, and 4) (card visual angle approximately $5.72° \times 5.72°$), with 4 target cards. To ensure the certainty of classification rules, test cards were required to share no more than one attribute with each target card, resulting in 24 test cards being selected (Lange et al., 2017). Four target cards and one test card were presented on the screen, with target cards presented directly above the test card, numbered 1–4 from left to right. Children were required to match target cards and test cards according to one of three possible classification rules and make key press responses according to the numbers corresponding to target cards. In each trial, an 800 ms fixation point was first presented at the center of the screen, followed by the stimulus interface, which disappeared upon key press; after an 800 ms interval, feedback was presented for 400 ms, informing participants whether the classification rule was correct (while also indicating whether to continue using the classification rule from the previous trial or find a new classification rule for the next trial) (detailed procedure shown in Figure 1D). Based on feedback, trials could be divided into switch trials and maintain trials. To reduce children's anticipation, rules were changed unpredictably, i.e., when the number of consecutive correct maintain trials was 4–6, a new rule was randomly introduced. The three classification rules were presented randomly in random order twice, with children completing at least 24 trials and at most 36 trials. The task lasted approximately 10 minutes. The accuracy and reaction time of switch trials and maintain trials were recorded, and children's total classification accuracy was calculated as the behavioral indicator of cognitive flexibility.

TMT. This test is a timed test that can assess children's cognitive flexibility ability and consists of two parts. In Part A, circles numbered 1–25 were included, requiring children to start from circle number 1 and connect all circles in ascending order by clicking the mouse (e.g., 1–2–3–…). In Part B, 25 circles were also included, but half were numbered with digits 1–13 and half with letters A–L, requiring participants to connect all circles in ascending, alternating order by drawing lines (e.g., 1–A–2–B–3–C–…). Part A lasted 2.5 minutes, and Part B lasted 5 minutes, with the actual time required to complete each part recorded. The time difference between Part B and Part A (i.e., switch cost) was calculated, with smaller differences indicating better cognitive flexibility.

It should be noted that in the 1-back, Go/Nogo, and Oddball tasks, correct responses to target stimuli were defined as hits, and incorrect responses to non-target stimuli were defined as false alarms. In the three tasks, after excluding invalid responses or trials with RT $<$ 200 ms in the formal experiment, the effective rate needed to reach 80% or above; trials with correct responses to non-target stimuli needed to account for more than 50% of effective trials. Data from children who did not meet these two conditions were treated as missing (Huizinga et al., 2006). In this study, the test-retest reliability ICCs for all task indicators in the EF test battery were: 1-back: 0.56; Digit Span: 0.44; Oddball: 0.21; Go/Nogo: 0.13; WCST: 0.35; TMT: 0.18.

2.2.2 Intelligence

Raven's Standard Progressive Matrices (SPM), developed by Raven et al. (1983) and revised by Zhang Houcan and Wang Xiaoping (1989), was used to assess children's observation ability and clear thinking ability. This test is a purely non-verbal intelligence test widely used for cross-cultural intelligence/reasoning ability assessment, evaluating individuals' non-verbal intelligence. The test consists of 5 units with 60 pictures, each being a progressive matrix composition. The matrix structure becomes increasingly complex, evolving from one level to multiple levels, and the required thinking operations also progress from direct observation to indirect abstract reasoning. One point is awarded for each correct answer, with a total score of 60. Based on children's age and norms, raw scores were converted to standard grade scores, with those at grade five considered intellectually deficient and excluded.

2.2.3 Demographic Variables

Children's gender, age, ethnicity, and family socioeconomic status (mainly father/mother education level) were reported by parents, and children's handedness was self-reported.

2.3 Research Procedure

Children were recruited from three public primary schools to participate in "cool" executive function assessments. Before data collection, informed consent was obtained from schools and head teachers, after which children were given a parent questionnaire containing informed consent information. Parents voluntarily signed whether they agreed to let their children participate in the study and completed related parent surveys. In the school computer room, four trained psychology graduate students served as research assistants responsible for data collection, which lasted 60 minutes. During the computer task testing phase, each child completed the executive function test battery, which included six tasks. For each task, research assistants guided children to understand task objectives and conducted practice trials. In each task, a practice trial correct rate of 90% or above was required to proceed to the formal experiment. To thank children for their participation, children who completed tasks each time received a beautiful school supply item.

2.4 Data Management and Analysis

IBM SPSS 26.0 (IBM Corp., Armonk, N.Y., USA) was used for double entry of questionnaire data, and Excel 2010's conditional formatting function was used to verify data accuracy. For computer task data, Excel 2010's classification summary function and vlookup function were used to calculate each child's indicators for each task. IBM SPSS 26.0, JASP 0.16.4.0 (JASP Team, 2022), Mplus 8.3 (Muthén & Muthén, 1998–2017, CA), and R 4.1.2 (R Core Team, 2020) were used to organize and analyze all data, including the following processes:

First, in IBM SPSS 26.0, data from the computer task testing phase were cleaned and transformed. For ease of understanding, data were transformed so that higher scores indicated better performance. Therefore, for switch cost in the TMT task, this study multiplied it by -1 to convert it to a negative value (i.e., –RT) (Brydges et al., 2014; Feng et al., 2022). The Expectation-Maximization Algorithm (EM) was used to impute missing data for each time point separately (Song et al., 2022).

Second, in JASP 0.16.4.0, descriptive statistics and correlation analysis were conducted to provide general characteristics of EF and math anxiety. In Mplus 8.3, 12 latent factor structures of EF were constructed and validated through confirmatory factor analysis, with models estimated using Maximum Likelihood (ML). Model fit indices included Chi-Square Test (^{2}), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root-Mean Square Error of Approximation (RMSEA), and Standardized Root-Mean-Square Residual (SRMR). Model fit criteria were as follows: CFI and TLI greater than 0.90, RMSEA less than 0.05, SRMR less than 0.08 (Hu & Bentler, 1999). By comparing AIC and BIC indices of well-fitting models, the model that best fit the behavioral data was selected as the structure of EF (Qiu & Lin, 2019).

Third, in R 4.1.2, a Gaussian Graphical Model (GGM) of partial correlation coefficients between the six EF tasks was estimated (Epskamp, 2015). The Least Absolute Shrinkage and Selection Operator (LASSO) was applied for regularization to carefully identify network edges, thereby accurately revealing the underlying network. To visually compare different networks, the network layout was unified using the average positions of nodes in the two networks, and the smallest edge value in the network was set to 0.04. Additionally, the EGAnet package (Younger et al., 2023) was further used to detect whether different tasks formed one or more communities within the network.

Finally, in Mplus 8.3, longitudinal invariance of the unity and diversity of EF was further examined.

3.1 Preliminary Analysis

First, descriptive statistics and correlation analysis were conducted on EF task-related indicators (see Table 1). Results showed that EF-related task indicators in this study were normally or approximately normally distributed ($-0.62 \leq$ Skewness ≤ 0.53). There were significant positive correlations between EF task performances at the same time point (T1: rs ≥ 0.10, ps < 0.05; T3: rs ≥ 0.09, ps < 0.05). There were also significant positive correlations between the same EF tasks across different time points (rs ≥ 0.17, ps < 0.001).

Table 1 Descriptive Statistics and Correlation Results of Variables

Note: 1-back and Digital Span are working memory-related tasks; Go/Nogo and Oddball are inhibitory control-related tasks; TMT and WCST are cognitive flexibility-related tasks. Red indicates positive correlation, blue indicates negative correlation, with darker colors indicating stronger correlations. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.2 Candidate Models of EF Structure at Different Time Points

To explore the structure of EF in middle childhood children, confirmatory factor analysis was used to construct 12 candidate models (also called competing models, including correlated factor models and bifactor models) for EF at two time points. Model fit results are shown in Table 2 . In correlated factor models, the "G," "I + W/C," "I/C + W," and "I + W + C" models, as well as the "G + W," "G + C + W," "G + I + W," and "G + I + C" models in bifactor models, all showed good fit across all time points, while the remaining models showed poor fit or failed to converge at least one time point. However, in bifactor models, some task indicators in the well-fitting models had non-significant loadings (ps > 0.05) on the inhibitory control-specific factor, working memory-specific factor, or cognitive flexibility-specific factor at both time points. Therefore, the "G," "I/W + C," "I + W/C," "I/C + W," and "I + W + C" models from correlated factor models were retained. Consistent with earlier research (Friedman et al., 2008; Miyake et al., 2000), the "I + W + C" model indicates the diversity of EF, with three factors showing high correlations (ps < 0.001). The "G" model indicates the unity of EF. Additionally, the "I/W + C," "I + W/C," and "I/C

+ W" models also showed good fit, supporting the view that inhibitory control, cognitive flexibility, and working memory are not independent of each other.

Table 2 Fit Indices of Correlated Factor Candidate Models of EF at Two Time Points

Note: In correlated factor models, W = working memory-related factor, I = inhibitory control-related factor, C = cognitive flexibility-related factor, G = general factor, I/C = common factor of inhibitory control and cognitive flexibility, I/W = common factor of inhibitory control and working memory, W/C = common factor of working memory and cognitive flexibility.

Five well-fitting competing correlated factor models were compared to determine the best-fitting model for children at different stages at each time point (results shown in Table 3). Model comparisons showed that at both time points, compared with the "G," "I/W + C," "I/C + W," and "I + W + C" models, the "I + W/C" model had smaller AIC and BIC values, indicating that the structure of EF in middle childhood children mainly exists in the form of a two-factor "I + W/C" model.

Table 3 Comparison of Well-Fitting Correlated Factor Models

3.3 Network Characteristics of EF-Related Tasks at Different Time Points

Next, network analysis was conducted on the six EF task indicators at two time points to understand how different tasks are associated. Results are shown in Figure 2 [Figure 2: see original paper]. Overall, the network models at the two time points were similar, with the network at T2 being more complex than at T1. Regarding network communities, the Spinglass algorithm in the EGAnet package showed only one community at both time points. This result emphasizes that EF outcomes change little in third to fourth grade children, representing a single component, but the relationships between these tasks gradually become more complex, meaning that EF continues to improve over time. Therefore, it is necessary to emphasize the unity of EF at this stage.

Note: Lines between nodes represent partial correlations between tasks, with thicker lines indicating larger partial correlation coefficients. To visually compare different networks, network layouts were unified using the average positions of nodes in the two networks. The colored areas in the rings around nodes depict predictability.

3.4 Longitudinal Invariance Testing of Unity and Diversity of EF

Based on the above results from latent variable analysis and network analysis, this study suggests that both unity and diversity of EF in middle childhood children cannot be ignored. However, bifactor models showed poor fit in this study. To measure the unity and diversity of EF, the general factor from the single-factor "G" model was extracted as an indicator of EF unity, and the component-related factors from the best-fitting two-factor "I + W/C" model were extracted as indicators of EF diversity. Factor loadings of unity and diversity of EF at the two time points are shown in Figure 3 [Figure 3: see original paper].

Note: The left panel shows the "G" model, where FaG is the general factor in the "G" model; the right panel shows the "I + W/C" model, where FaI and FaWC are component factors in the "I + W/C" model, corresponding to the inhibitory control factor and the working memory/cognitive flexibility common factor, respectively. Standardized coefficients to the left of the slash are from T1, and those to the right are from T2; standard errors are in parentheses. *** $p < 0.001$.

Multiple Groups Confirmatory Factor Analysis was used to construct a series of nested models to test longitudinal invariance of EF unity and diversity and math anxiety. When full invariance could not be satisfied, the Forward method (Jung & Yoon, 2016) was used for partial invariance settings. Model fit criteria were as follows: CFI and TLI greater than 0.90, RMSEA less than 0.05, SRMR less than 0.08 (Hu & Bentler, 1999). Nested model comparisons were evaluated based on changes in fit indices $\Delta\ ^2$, ΔCFI, and ΔRMSEA, i.e., when $\Delta\ ^2$ did not reach significance, it indicated that the more restricted model did not result in worse model fit, and invariance was acceptable; when the change in CFI was less than 0.010 and the change in RSMEA was less than 0.015, it indicated that the more restricted model did not result in worse model fit, and invariance was acceptable (Chen, 2007; Satorra & Bentler, 2001). Since chi-square tests are greatly affected by sample size (Feng et al., 2022), the latter criteria were relied upon more heavily for judgment. Longitudinal invariance test results (shown in Table 4) indicated that both EF unity (i.e., the general factor in the "G" model) and diversity (i.e., the component-related factors in the "I + W/C" model) satisfied full metric invariance.

Table 4 Longitudinal Invariance Tests of Unity and Diversity of EF

Note: G = general factor, I = inhibitory control-related factor, W/C = working memory and cognitive flexibility common factor.

4.1 Optimal Structure of EF in Middle Childhood Children

The structure of individual EF changes with age (Karr et al., 2018). So what form does EF take in middle childhood children in the Chinese cultural context? The latent variable analysis at two time points in this study found that single-factor ("G" model), two-factor ("I + W/C" model, "C + I/W" model, and "I/C + W" model), and three-factor ("I + W + C" model) models all showed good fit with behavioral data. Further model comparisons showed that the two-factor "I + W/C" model was the best-fitting model for middle childhood, indicating that EF at this stage exists in the form of inhibitory control and a common factor of working memory and cognitive flexibility. This means that the inhibitory control factor has been effectively separated at this time, but working memory and cognitive flexibility share common components and cannot be separated (Klauer et al., 2010). Earlier studies found that in normally developing children around age 9, specific executive functions could not be distinguished from each other; later-developing components such as cognitive flexibility could not be separated until ages 10–11 (Brydges et al., 2014). This study confirms the general trend observed in previous research, with results also showing that working memory and cognitive flexibility share common components that are difficult to separate at this stage. Possible reasons are: on the one hand, cognitive flexibility has higher correlations with working memory; on the other hand, compared with inhibitory control, both appear relatively later and therefore may be difficult to separate.

Additionally, this study used the 1-back task and Digital Span task to measure working memory. Children needed not only to simply remember currently presented information but also to continuously extract previous information to achieve better performance, involving the updating process of memory maintenance and retrieval. Similarly, cognitive flexibility tasks also require updating processes. This study used a simplified WCST task and TMT task to measure cognitive flexibility, both of which also involve memory maintenance, retrieval, and updating processes. Therefore, both working memory and cognitive flexibility essentially require "retrieving" previous information and "maintaining" current information in the mind, which may be a shared process between the two components—the working memory/cognitive flexibility common factor.

4.2 Indicators for Measuring "Unity and Diversity" of EF in Middle Childhood Children

Regarding the operational definition of EF in childhood, most researchers extract a single component using multiple tasks, treating it as a single structure (e.g., Blair et al., 2011; Clark et al., 2013; Sulik et al., 2015; Vrantsidis et al., 2019), or extract multiple independent components/latent variables using multiple tasks (e.g., Matte-Gagné et al., 2018; Simanowski & Krajewski, 2019), exploring unity or diversity separately. Teuber (1972) first proposed the unity and diversity of EF in the article "Unity and Diversity of Frontal Lobe Functions," which has been validated in individual difference studies of EF. In this study, latent variable analysis showed that the "I + W/C" model was the optimal model for EF in middle childhood children; however, network analysis showed that the six EF task indicators were identified as one community at both time points, highlighting the unity of EF structure at this stage. The two analysis results complement each other, emphasizing that unity and diversity are not mutually exclusive and both play important roles in child development, and that EF research needs to examine both unity and diversity simultaneously.

Initially, the unity and diversity of EF were assessed through correlations (Miyake et al., 2000). In correlated factor models, unity and diversity are reflected in factor correlations. That is, when correlations between factors are greater than 0, it indicates the existence of unity; when they are less than 1.0, it indicates the existence of diversity. In this study, although the six tasks assessing EF showed significant correlations, the correlations between each pair were not consistently low. Tasks assessing the same EF component showed strong, significant correlations with each other; tasks assessing different EF components showed less strong correlations with each other, showing some signs of convergent and discriminant validity. Additionally, whether the single-factor "G" model, two-factor "I/W + C" model, "I + W/C" model, and "I/C + W" model, or the three-factor "I + W + C" model, all showed good fit with behavioral data. Correlations between component-related factors did not reach 1.00, consistent with existing research findings (Feng et al., 2022; Friedman & Miyake, 2017; Miyake et al., 2000, 2001), indicating that EF has both unity and diversity.

Cognitive skill development has continuity and hierarchical nature, with higher-level cognitive skills directly related to EF developing based on simpler cognitive skills (Garon et al., 2008). These simpler cognitive skills are necessary conditions for effectively performing various EF tasks. Therefore, it is necessary to simultaneously focus on the role of simpler cognitive skills (i.e., "unity") and cognitive skills directly related to EF components (i.e., "diversity"). Friedman et al. (2008) constructed a bifactor model to estimate an alternative parameter to correlation—latent variables, where the common latent variable of EF in the model indicates unity, and specific factors reflect diversity. However, Friedman et al. did not consider age differences in EF structure. Related research (Brydges et al., 2014; a meta-analysis see Karr et al., 2018) and this study consistently found that bifactor models show poor fit in childhood, which may mean that individual EF is still in the developmental stage in middle childhood and has not yet differentiated into specific component-related factor structures. Some researchers have used factor analysis methods to extract different components from multiple tasks as indicators of unity and diversity (Friedman et al., 2011; Huizinga et al., 2006; Lehto et al., 2003; Miyake et al., 2000, 2001; Wiebe et al., 2008). Therefore, based on results from latent variable analysis and network analysis, this study selected the general factor from the "G" model as the indicator of unity, and the component-related factors from the best-fitting "I + W/C" model as indicators of diversity.

4.3 Stability of Unity and Diversity of EF in Middle Childhood Children

In existing research exploring EF structure, most are cross-sectional designs, and the few studies focusing on EF development only recruit multiple age groups for comparison (e.g., Davidson et al., 2006; Ferguson et al., 2021; Hartung et al., 2020; Huizinga et al., 2006; Laureys et al., 2022; Lee et al., 2013; Lehto et al., 2003; Xu et al., 2013). To our knowledge, only four studies have used longitudinal designs and tested longitudinal invariance to verify the stability of EF-related structures. Usai et al. (2014) tracked EF in 5–6-year-old children (N = 145) at one-year intervals, finding that the two-factor "I + W/C" model showed good fit at both time points, and longitudinal invariance testing satisfied partial scalar invariance. Brydges et al. (2014) focused on the differentiation of EF in middle to late childhood, tracking 8–9-year-old children (N = 135) for two years. Longitudinal invariance testing only satisfied configural invariance, indicating that factor structure changed during the investigation period, with single-factor models showing good fit at T1 and "W + I/C" models showing good fit at T2. However, both studies had small sample sizes, which is not favorable for longitudinal invariance testing. Friedman et al. (2016) tracked 786 same-sex twins from late adolescence to early adulthood, finding that the three-factor model in correlated factor models (i.e., "I + W + C" model) and the "C + U + S" model in bifactor models (corresponding to the "G + W + C" model in this study) showed good fit, but only the three-factor model in correlated factor models satisfied metric invariance. Blair et al. (2014) tracked EF in 1,292 36-month-old children, finding that the single-factor model (i.e., "G" model) satisfied partial metric invariance. These four studies examined different age stages; only one study satisfied partial scalar invariance, which cannot provide strong evidence for the stability of EF structure at each stage; they only focused on one aspect of unity or diversity, not simultaneously examining both unity and diversity from an integrated perspective of EF. This study used multiple-group confirmatory factor analysis to test longitudinal invariance of the "G" model and the best-fitting "I + W/C" model in correlated factor models, finding that both the "I + W/C" model and the "G" model satisfied full metric invariance, indicating that the above structures did not change during the testing period and were relatively stable, consistent with current two-time-point studies (Brydges et al., 2014; Usai et al., 2014). This also suggests the stability of EF unity and diversity at this stage. This study is the first domestic study to test the longitudinal invariance of EF unity and diversity from two time points, not only revealing the structure of EF in middle childhood children but also providing reliable evidence for this structure; it also emphasizes the necessity of simultaneously examining both unity and diversity of EF in middle childhood children.

It is worth noting that there is still debate among researchers about the meaning of the general factor of EF. The bifactor "G + C + W" model has been confirmed as the best model for EF in many studies (e.g., Feng et al., 2022; Friedman et al., 2011, 2016; Glisky et al., 2021; Gustavson et al., 2015; Valian, 2015), but this model cannot separate the inhibitory control-specific factor. Many EF tasks require some form of inhibition (e.g., inhibiting responses, distraction, memory representations), leading some researchers to believe that the EF general factor is the inhibitory control factor (e.g., Valian, 2015). However, this concept of inhibition may be too broad, grouping together processes that are conceptually and empirically separable (Friedman & Miyake, 2004, 2017). Friedman and Miyake (2017) argue that the EF general factor reflects individual differences in the ability to maintain and manage goals, which are general requirements for all EF tasks. Individuals use goals to correct ongoing tasks and maintain relevant information, a view that is gaining increasing recognition in research (e.g., Feng et al., 2022).

Undeniably, this study's results still have certain limitations for understanding the development of EF structure. On the one hand, based on the EF unity/diversity framework, this study selected related tasks to assess three common components of EF, and results may be influenced by EF components and related tasks. Further understanding of the composition factors of EF structure and determination of how to measure the specificity of these components is needed. On the other hand, this study only examined the structure of EF in middle childhood children (third to fourth grade), which to some extent limits the global depiction of developmental trends in EF structure. To better understand the structural characteristics of EF at this stage, more representative samples need to be selected for larger-scale validation work to further extend the results of the current study.

---

5 Research Conclusion

In the context of Chinese culture, EF in middle childhood children exists in the form of inhibitory control and a cognitive flexibility/working memory common factor (i.e., the "I + W/C" model), and the unity (general factor) and diversity (component-related factors) of EF are relatively stable.

References

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Blair, C., Protzko, J., & Ursache, A. (2011). Self-regulation and the development of early literacy. In: D. Dickinson, S. Neuman (Eds.), *Handbook of early literacy research* Vol. 3 (pp. 20–35). New York: Guilford Press.

Blair, C., Raver, C. C., Berry, D. J., & The Family Life Project Investigators. (2014). Two approaches to estimating the effect of parenting on the development of executive function in early childhood. *Developmental Psychology, 50*(2), 554–565. https://doi.org/10.1037/a0033647

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence, 47*, 34–43. https://doi.org/10.1016/j.intell.2014.08.010

Cai, Y., Dong, S., Yuan, S., & Hu, C. (2020). Network analysis and its applications in psychology. *Advances in Psychological Science, 28*(01), 178–195. [蔡玉清, 董书阳, 袁帅, 胡传鹏. (2020). 变量间的网络分析模型及其应用. 心理科学进展, 28(01), 178–195.]

Carlson, S. M., White, R. E., & Davis-Unger, A. C. (2014). Evidence for a relation between executive function and pretense representation in preschool children. *Cognitive Development, 29*, 1–16. https://doi.org/10.1016/j.cogdev.2013.09.001

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Clark, C. A. C., Sheffield, T. D., Wiebe, S. A., & Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Development, 84*(2), 662–677. https://doi.org/10.1111/j.1467-8624.2012.01854.x

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the Cantab. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology, 25*(2), 242–254. https://doi.org/10.1076/jcen.25.2.242.13639

Downes, M., Bathelt, J., & De Haan, M. (2017). Event-related potential measures of executive functioning from preschool to adolescence. *Developmental Medicine & Child Neurology, 59*(6), 581–590. https://doi.org/10.1111/dmcn.13395

Epskamp, S. (2015). bootnet: Bootstrap methods for various network estimation routines. R-Package.

Espy, K. A., Kaufmann, P. M., McDiarmid, M. D., & Glisky, M. L. (1999). Executive functioning in preschool children: Performance on A-not-B and other delayed response format tasks. *Brain and Cognition, 41*(2), 178–199. https://doi.org/10.1006/brcg.1999.1117

Feng, J., Zhang, L., Chen, C., Sheng, J., Ye, Z., Feng, K., Liu, J., Cai, Y., Zhu, B., Yu, Z., Chen, C., Dong, Q., & Xue, G. (2022). A cognitive neurogenetic approach to uncovering the structure of executive functions. *Nature Communications, 13*, Article 4588. https://doi.org/10.1038/s41467-022-32383-0

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports, 11*, Article 1382. https://doi.org/10.1038/s41598-020-80866-1

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology, 52*(2), 326–340. https://doi.org/10.1037/dev0000075

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology, 47*(5), 1410–1430. https://doi.org/10.1037/a0023750

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Glisky, E. L., Alexander, G. E., Hou, M., Kawa, K., Woolverton, C. B., Zigman, E. K., Nguyen, L. A., Haws, K., Figueredo, A. J., & Ryan, L. (2021). Differences between young and older adults in unity and diversity of executive functions. *Aging, Neuropsychology, and Cognition, 28*(6), 829–854. https://doi.org/10.1080/13825585.2020.1830936

Goldstein, S., & Naglieri, J. A. (2014). *Handbook of executive functioning.* New York: Springer.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Gustavson, D. E., Miyake, A., Hewitt, J. K., & Friedman, N. P. (2015). Understanding the cognitive and genetic underpinnings of procrastination: Evidence for shared genetic influences with goal management and executive function abilities. *Journal of Experimental Psychology: General, 144*(6), 1063–1079. https://doi.org/10.1037/xge0000110

Hartung, J., Engelhardt, L. E., Thibodeaux, M. L., Harden, K. P., & Tucker-Drob, E. M. (2020). Developmental transformations in the structure of executive functions. *Journal of Experimental Child Psychology, 189*, Article 104681. https://doi.org/10.1016/j.jecp.2019.104681

Hu, L.—T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hughes, C., & Devine, R. T. (2019). For better or for worse? Positive and negative parental influences on young children's executive function, *Child Development, 90*(2), 593–609. https://doi.org/10.1111/cdev.12915

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*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

Isquith, P. K., Roth, R. M., & Gioia, G. A. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child, 2*(2), 125–132. https://doi.org/10.1080/21622965.2013.748389

JASP Team. (2022). JASP (Version 0.16.4) [Computer software]. Retrieved from https://jasp—stats.org/.

Jung, E., & Yoon, M. (2016). Comparisons of three empirical methods for partial factorial invariance: Forward, backward, and factor—ratio tests. *Structural Equation Modeling, 23*(4), 567–584. https://doi.org/10.1080/10705511.2015.1138092

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin, 144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Klauer, K. C., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2010). Understanding the role of executive control in the implicit association test: Why flexible people have small IAT effects. *Quarterly Journal of Experimental Psychology, 63*(3), 595–619. https://doi.org/10.1080/17470210903076826

Lange, F., Seer, C., Müller-Vahl, K., & Kopp, B. (2017). Cognitive flexibility and its electrophysiological correlates in Gilles de la Tourette syndrome. *Developmental Cognitive Neuroscience, 27*, 78–90. https://doi.org/10.1016/j.dcn.2017.08.008

Laureys, F., De Waelle, S., Barendse, M. T., Lenoir, M., & Deconinck, F. J. (2022). The factor structure of executive function in childhood and adolescence. *Intelligence, 90*, Article 101600. https://doi.org/10.1016/j.intell.2021.101600

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development, 84*(6), 1933–1953. https://doi.org/10.1111/cdev.12096

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*(1), 59–80. https://doi.org/10.1348/026151003321164627

Maasalo, K., Lindblom, J., Kiviruusu, O., Santalahti, P., & Aronen, E. T. (2021). Longitudinal associations between inhibitory control and externalizing and internalizing symptoms in school-aged children. *Development and Psychopathology, 33*(3), 843–855. https://doi.org/10.1017/s0954579420000176

Matte-Gagné, C., Bernier, A., Sirois, M.-S., Lalonde, G., & Hertz, S. (2018). Attachment security and developmental patterns of growth in executive functioning during early elementary school. *Child Development, 89*(3), e167–e182. https://doi.org/10.1111/cdev.12807

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society, 16*(3), 495–505. https://doi.org/10.1017/S1355617710000093

McCoy, D. C. (2019). Measuring young children's executive function and self-regulation in classrooms and other real-world settings. *Clinical Child and Family Psychology Review, 22*(1), 63–74. https://doi.org/10.1007/s10567-019-00285-1

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

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*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130*(4), 621–640. https://doi.org/10.1037/0096-3445.130.4.621

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.

Pelegrina, S., Molina, R., Rodríguez-Martínez, E. I., Linares, R., & Gómez, C. M. (2020). Age-related changes in selection, recognition, updating and maintenance information in WM. An ERP study in children and adolescents. *Biological Psychology, 157*, Article 107977. https://doi.org/10.1016/j.biopsycho.2020.107977

Qi, Y., Yang, G., Fu, D., Li, Z., & Liu, X. (2021). Developmental neuroscience of cognitive control: The future path and layout. *Science Sinica Vitae, 51*(6), 634–646. [齐玥, 杨国春, 付迪, 李政汉, 刘勋. (2021). 认知控制发展神经科学：未来路径与布局. 中国科学生命科学, 51(6), 634–646.]

Qiu, Z., & Lin, B. (2019). *Principle and application of the structural equation model* (second edition). Beijing: Chinese light industry press. [邱皓政, 林碧芳. (2019). 结构方程模型的原理与应用（第二版）. 北京：中国轻工业出版社.]

R Core Team (2020). *A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing.

Raven, J. C., Court, J. H., & Raven J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales* (Section 4: Advanced Progressive Matrices). London: H. K. Lewis.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507–514. https://doi.org/10.1007/BF02296192

Seer, C., Lange, F., Georgiev, D., Jahanshahi, M., & Kopp, B. (2016). Event-related potentials and cognition in Parkinson's disease: An integrative review. *Neuroscience and Biobehavioral Reviews, 71*, 691–714. https://doi.org/10.1016/j.neubiorev.2016.08.003

Shi, R., Sharpe, L., & Abbott, M. (2019). A meta-analysis of the relationship between anxiety and attentional control. *Clinical Psychology Review, 72*, Article 101754. https://doi.org/10.1016/j.cpr.2019.101754

Simanowski, S., & Krajewski, K. (2019). Specific preschool executive functions predict unique aspects of mathematics development: A 3-year longitudinal study. *Child Development, 90*(2), 544–561. https://doi.org/10.1111/cdev.12909

Song, Z., Guo, L., & Zheng, T. (2022). Comparison of missing data handling methods in cognitive diagnosis: Zero replacement, multiple imputation and maximum likelihood estimation. *Acta Psychologica Sinica, 54*(4), 426—440. [宋枝璘, 郭磊, 郑天鹏. (2022). 认知诊断缺失数据处理方法的比较：零替换、多重插补与极大似然估计法. 心理学报, 54(4), 426—440.]

Strauss, E., Sherman, E. M. X., & Spreen, O. (2006). *A compendium of neuropsychological tests* (3rd edn). New York: Oxford University Press.

Sulik, M. J., Blair, C., Mills-Koonce, R., Berry, D., Greenberg, M., & The Family Life Project Investigators. (2015). Early parenting and the development of externalizing behavior problems: Longitudinal mediation through children's executive function. *Child Development, 86*(5), 1588–1603. https://doi.org/10.1111/cdev.12386

Takacs, Z. K., & Kassai, R. (2019). The efficacy of different interventions to foster children's executive function skills: A series of meta-analyses. *Psychological Bulletin, 145*(7), 653–697. https://doi.org/10.1037/bul0000195

Teuber, H. L. (1972). Unity and diversity of frontal lobe functions. *Acta Neurobiologiae Experimentalis, 32*(2), 615–656.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, 54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Usai, M. C., Viterbori, P., Traverso, L., & De Franchis, V. (2014). Latent structure of executive function in five- and six-year-old children: A longitudinal study. *European Journal of Developmental Psychology, 11*(4), 447–462. https://doi.org/10.1080/17405629.2013.840578

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*, 3–24.

Vrantsidis, D., Clark, C. A. C., Chevalier, N., Espy, K. A., & Wiebe, S. (2019) Socioeconomic status and executive function in early childhood: Exploring proximal mechanisms, *Developmental Science, 23*, Article 12917. https://doi.org/10.1111/desc.12917

Weschler. (2002). *Wechsler Preschool and Primary Scale of Intelligence - Third Edition.* Harcourt Assessment.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology, 44*(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575

Wild, K. V., & Musser, E. D. (2014). The Cambridge neuropsychological test automated battery in the assessment of executive functioning. In: S. Goldstein and J. A. Naglieri (eds.), *Handbook of executive function* (pp.171–190). New York: Springer.

Xu, F., Han, Y., Sabbagh, M. A., Wang, T., Ren, X., & Li, C. (2013). Developmental differences in the structure of executive function in middle childhood and adolescence. *PLoS ONE, 8*(10), Article 77770. https://doi.org/10.1371/journal.pone.0077770

Yaple, Z., & Arsalidou, M. (2018). N-back working memory task: Meta-analysis of normative fMRI studies with children. *Child Development, 89*(6), 2010–2022. https://doi.org/10.1111/cdev.13080

Younger, J. W., O'Laughlin, K. D., Anguera, J. A., Bunge, S. A., Ferrer, E. E., Hoeft, F., McCandliss, B. D., Mishra, J., Rosenberg-Lee, M., Gazzaley, A., & Uncapher, M. R. (2023). Better together: Novel methods for measuring and modeling development of executive function diversity while accounting for unity. *Frontiers in Human Neuroscience, 17*, Article 1195013. https://doi.org/10.3389/fnhum.2023.1195013

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives, 6*(4), 354–360. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zhang, H., & Wang, X. (1989). Standardization research on Raven's standard processive matrices in China. *Acta Psychologica Sinica, 21*(2), 113–121. [张厚粲, 王晓平. (1989). 瑞文标准推理测验在我国的修订. 心理学报, 21(2), 113–121.]

Corresponding Author Email: Jiwei Si (sijiwei@126.com)

Author Contributions Statement:

Bijuan Huang, Jiwei Si: Conceived research ideas and designed research protocol;
Bijuan Huang, Hongxiang Zhu, Chang Liu, Hongmin Feng, Ruifan Luo: Conducted experiments;
Bijuan Huang, Hongxia Li: Collected, cleaned, and analyzed data;
Bijuan Huang: Drafted manuscript;
Bijuan Huang, Jiwei Si: Revised final version of manuscript.