DOI: 10.1111/cdev.13888

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

Development of complex executive function over childhood: Longitudinal growth curve modeling of performance on the Groton Maze Learning Task

Thomas B. McGuckian[1,2] | Peter H. Wilson[1,2] | Rich D. Johnston[2,3,4] | Shahin Rahimi-Golkhandan[5] | Jan Piek[6] | Dido Green[7,8] | Jeffrey M. Rogers[9] | Paul Maruff[10] | Bert Steenbergen[11] | Scott Ruddock[12]

[1]Healthy Brain and Mind Research Centre, Australian Catholic University, Melbourne, Victoria, Australia
[2]School of Behavioural and Health Sciences, Australian Catholic University, Brisbane, Queensland, Australia
[3]Sports Performance, Recovery, Injury and New Technologies (SPRINT) Research Centre, Australian Catholic University, Brisbane, Queensland, Australia
[4]Carnegie Applied Rugby Research Centre, Institute for Sport, Physical Activity and Leisure, Leeds Beckett University, Leeds, UK
[5]School of Psychology and Wellbeing, University of Southern Queensland, Toowoomba, Queensland, Australia
[6]Curtin University, Perth, Western Australia, Australia
[7]Jönköping University, Jönköping, Sweden
[8]Brunel University, London, UK
[9]University of Sydney, Sydney, New South Wales, Australia
[10]University of Melbourne, Melbourne, Victoria, Australia
[11]Radboud University, Nijmegen, The Netherlands
[12]La Trobe University, Melbourne, Victoria, Australia

Correspondence
Thomas B. McGuckian, Healthy Brain and Mind Research Centre (HBMRC) and School of Behavioural and Health Sciences, Australian Catholic University, Melbourne, VIC, Australia.
Email: thomas.mcguckian@acu.edu.au

Funding information
Australian Research Council, Grant/Award Number: DP1094535

Abstract
This longitudinal study modeled children's complex executive function (EF) development using the Groton Maze Learning Task (GMLT). Using a cohort-sequential design, 147 children (61 males, 5.5–11 years) were recruited from six multicultural primary schools in Melbourne and Perth, Australia. Race/ethnicity data were not available. Children were assessed on the GMLT at 6-month intervals over 2-years between 2010 and 2012. Growth curve models describe age-related change from 5.5 to 12.5 years old. Results showed a quadratic growth trajectory on each measure of error—that is, those that reflect visuospatial memory, executive control (or the ability to apply rules for action), and complex EF. The ability to apply rules for action, while a rate-limiting factor in complex EF, develops rapidly over early-to-mid childhood.

Theoretical and empirical models consider executive function (EF) to be hierarchical in structure, consisting of (i) a domain-general module termed *cognitive control*, which is concerned with coordinating cognitive operations, implementing strategies for problem solving, and monitoring errors (Thomas et al., 2013) and (ii) a subset of more specialized (modality-specific) operations including working memory, inhibition, and executive attention (Diamond, 2013; Tirapu-Ustarroz et al., 2018). In this paper, we report on a longitudinal investigation

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; CSD, cohort sequential design; EF, executive function; GCM, growth curve modeling; GMLT, Groton Maze Learning Task; PFC, prefrontal cortex.

of EF over childhood using a single, well-validated maze learning paradigm, the Groton Maze Learning Task (GMLT). Distinct growth curve models are presented that describe age-related changes in higher-order cognitive control processes (specifically, rule-based error monitoring) and *visuospatial working memory*.

Development of complex EF in children

It is widely accepted that EF develops rapidly over childhood and into the adolescent period, in synergy with the heightened demands of learning, schooling, and sociocultural participation. The development of complex EF is, indeed, critical to the acquisition of academic and life skills (Zelazo & Carlson, 2017), and predicts academic achievement (Best et al., 2011; Pascual et al., 2019) and later adaptive function (Benson et al., 2013). These relations underscore the importance of understanding the pattern of growth in the individual cognitive processes that make up EF and for developing a framework for objectively evaluating EF in children who do not meet expected intellectual milestones.

Developmental studies reveal a period of ongoing development in EFs over childhood, adolescence and adulthood, with more complex aspects of cognitive control among the last to develop (Alloway & Gathercole, 2012; Casey et al., 2005). Of particular importance is the ability to apply *action-oriented rules* in the context of executive control and problem solving (Zelazo & Carlson, 2017); such rules govern the selection of a behavior from a set of possible actions, for example, in the DCCS, *if* a color game, *then if* red, place card here. In general, higher-order cognitive control—as shown by the use of complex rules—has a more prolonged period of development than low-level cognitive operations like working memory and draws more heavily on maturation of the late-developing prefrontal cortex (PFC; Botvinick & Cohen, 2014). Taken together, this suggests that maturational aspects of EF are best described by some type of growth curve. However, the comparison of growth trajectories for specific cognitive operations is difficult because different tests are often used at different age points (Mirabella, 2021). As well, longitudinal studies of EF often do not span the entire school-aged period, which further limits the modeling of developmental change (Zelazo & Carlson, 2020).

Behavioral and cognitive neuroscience studies show that performance of complex visuospatial tasks that involve rule-governed behavior (like maze learning) is dependent on separable cognitive processes that include planning based on prior behavior, feedback from errors, and holding information in memory about the outcomes of prior moves or trials (Carlson et al., 2019). Hidden pathway maze-learning paradigms have a long history in experimental psychology and neuropsychology (Thomas et al., 2014). These paradigms generally require participants to apply set rules about permissible moves (or touch responses) in order to find and learn a maze path that is hidden from view. The GMLT is a particularly well-validated, computer-controlled version of the Milner Hidden Pathway Maze learning test, used in the study reported here. In the GMLT (Figure 1), performers are instructed to follow a set of rules that govern the type of moves that are allowed in order to learn a hidden maze (within a square grid of tiles) over repeated trials. For the GMLT, the following four rules apply: (i) move one tile at a time; (ii) do not move diagonally; (iii) do not backtrack along the path; and (iv) return to the last correct tile after an error. After each move, visual feedback is temporarily provided on the selected tile (green tick or red cross) to indicate whether the performer has correctly hit on the hidden path or not. If they correctly hit the path, they can make their next selection. If they hit off the path, they are required to return to the last correct tile and select an alternate next tile. A trial is complete when the performer reaches the end of the maze.

While completing trials, two types of errors are recorded by the computer. *Rule-break errors*, which occur when a move is made that is not allowed by the

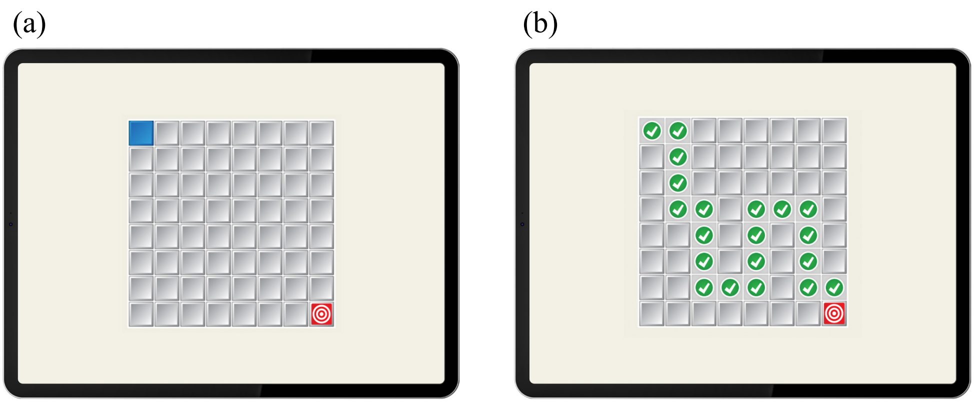

FIGURE 1 Example images of the Groton Maze Learning Task, showing (a) the task display during a trial, and (b) the task display after the entire path has been successfully completed.

four defined rules, reflect the ability to remember and apply set rules (i.e., error monitoring). *Legal errors* (also termed *spatial errors*), where responses are in accordance with the four defined rules but to locations that do not lie on the hidden pathway, reflect visuospatial memory (Thomas et al., 2013). *Total errors*, calculated as the sum of *rule-break errors* and *legal errors*, gives a composite measure of EF that includes aspects of visuospatial working memory, executive attention, and inhibition. *Response duration* is the total time taken to complete all trials, and reflects psychomotor speed and general performance efficiency (Thomas et al., 2011, 2016).

The neuropsychological literature provides strong evidence for the independence of cognitive operations measured by different GMLT metrics. *Double dissociations* in the performance profiles of neurological syndromes are particularly powerful in demonstrating this point. In patients with lesions of the PFC, we see a high proportion of rule-break errors, relative to spatial errors (P. J. Snyder et al., 2008), while patients with damage to the medial temporal lobe show the reverse pattern of performance with a higher proportion of spatial learning errors. Further, patients with schizophrenia see greater improvement for rule-break errors than spatial errors when treated with medication, while children treated for Attention Deficit Hyperactivity Disorder with stimulant medication show greater improvement for spatial errors than rule-break errors (A. M. Snyder et al., 2008).

Studies of convergent and divergent validity support the independence of rule-break and spatial errors (Pietrzak et al., 2008). For school-age children, Thomas et al. (2016) showed that rule-breaks on the GMLT correlated moderately (r = .50) with other measures of working memory that also involve error monitoring (like Continuous Paired Associate Learning), but not with simple measures of spatial updating (Corsi Blocks). As well, cross-sectional comparisons reveal that difficulties in error monitoring are evident in younger children (5–6 years of age) who commit proportionally more rule-breaks than children aged 9 years, but only on more complex 8×8 or 10×10 mazes (Thomas et al., 2011). This result suggests that the ability to enlist error monitoring strategies is not well developed in younger children, especially when the spatial difficulty of the task is high. Such errors are virtually absent by adulthood.

The primary aim of the study presented here was to examine the development over childhood of complex EF and its component processes using a well-validated paradigm and powerful (cohort-sequential) longitudinal design. Few longitudinal studies have afforded growth modeling across the entire primary-school aged period. Moreover, our use of the GMLT enabled repeat testing without the confound of learning effects (Snyder et al., 2005), and a componential analysis of cognitive operations. Specifically, we were interested in comparing the pattern of age-related change between the visuospatial working memory and rule-based error monitoring components of EF. We predicted that growth trajectories would differ between these measures: first, a linear growth trend was predicted on visuospatial working memory consistent with patterns observed on related measures like Corsi blocks (Thomas et al., 2016); second, a quadratic trend was predicted on the measure of rule-based error monitoring (an aspect of executive attention), based on the observation that higher-order EFs show rapid consolidation over the early-to-middle childhood period (Zelazo & Carlson, 2017). Despite the predictions outlined above, the current analyses are considered relatively exploratory given the somewhat novel approach and limited longitudinal studies of primary-school aged children on which the study is based.

METHOD

Participants

Our sample of typically developing children were recruited as part of a longitudinal study of motor and cognitive development in children (Ruddock et al., 2016). A total of 147 typically developing children (61 male) aged between 5.5 and 11 years of age were recruited from six mainstream (public and independent) primary schools in the greater Melbourne and Perth metropolitan areas. Specific racial and ethnic data were not collected originally, and thus were not available, but the sample reflected the multicultural communities of these areas. The Australian population has ancestries (of at least one parent) of approximately 57% European, 31% Oceanian (including 'Australian'), 17% Asian, 3% Indigenous, 3% North African and Middle Eastern, 1% Peoples of the Americas, 1% Sub-Saharan African (Australian Bureau of Statistics, 2021). All children were identified by parents on a developmental questionnaire as being free of any major medical or neurological condition, and none reported an intellectual disability. Demographics of the sample measured at baseline are provided in Table 1. Overall motor ability was assessed using the McCarron Assessment of Neuromuscular Development (McCarron, 1997); a Neurodevelopmental Index score of >80 (20th percentile) was set as the cut point to include typically developing children in the study at Time 0 (Hyde & Wilson, 2013).

Groton Maze Learning Task

Complex EF was assessed using an 8×8 version of the GMLT (Cogstate, 2018), presented centrally on a 12-inch touchscreen PC. Hidden beneath the grid of tiles is a 20-step pathway that leads from a start location at the top-left corner to an end location at the bottom-right

TABLE 1 Descriptive statistics of typically developing children at baseline

Age at baseline	Sample size		Mean age in years (SD)	
	Girls	Boys	Girls	Boys
5.5	3	—	5.81 (0.19)	—
6	6	—	6.31 (0.15)	—
6.5	11	9	6.70 (0.14)	6.63 (0.15)
7	5	5	7.18 (0.22)	7.13 (0.10)
7.5	10	6	7.72 (0.17)	7.75 (0.14)
8	6	11	8.19 (0.14)	8.22 (0.17)
8.5	7	4	8.68 (0.15)	8.75 (0.15)
9	9	6	9.28 (0.17)	9.24 (0.13)
9.5	7	6	9.76 (0.16)	9.69 (0.11)
10	8	8	10.23 (0.17)	10.23 (0.18)
10.5	9	3	10.67 (0.11)	10.75 (0.17)
11	5	3	11.37 (0.17)	11.56 (0.61)

Note: $N = 147$. Age at baseline was used to determine cohorts. Age was measured in months and converted to years. Example: a participant aged 6 years and 6 months = 6.5 years.

corner of the grid (see Figure 1 for an example). The child is instructed explicitly to solve the maze by finding the hidden pathway while adhering to four rules: (i) move one tile at a time; (ii) do not move diagonally; (iii) do not backtrack along the path; and (iv) do return to the last correct tile after an error. After each move, the computer provides temporary feedback on the selected tile (green tick or red cross) to indicate whether the child has correctly hit on the hidden path or not. If they correctly hit the path, they can make their next selection. If they hit off the path, they are required to return to the last correct tile and select an alternate next tile. A trial is complete when the child reaches the end location. A total of five trials were completed at each timepoint, and the correct path remained the same for each of the five trials. The goal is to find the pathway using the least moves possible and, hence, the fewest errors. A new path was used for each successive timepoint. Normative data for the 8×8 GMLT is only available for children aged 6–9 years: mean number of total errors in a healthy standardization sample is 52.9 (SD = 14.8) at 6 years of age, 50.0 (22.2) at 7 years, 43.1 (20.3) at 8 years, and 30.5 (12.3) at 9 years (Cogstate, 2018).

Four primary outcomes were calculated from the GMLT. *Rule-break errors* was calculated as the total number of times any of the four rules were broken, summed for all five trials. *Legal errors* was calculated as the total number of times a response was in accordance with the rules but to a location that did not lie on the hidden pathway, summed for all five trials. *Total errors* was calculated as the sum of *rule-break errors* and *legal errors*. *Response duration* was calculated as the total time (in seconds) taken to complete all five trials. (Thomas et al., 2013).

Procedure

Children were assessed every 6 months over a 2-year period from 2010 to 2012. Children were recruited in the first year, and followed for a maximum of 2 years, or until they no longer attended primary school. Ethics approval was obtained from the relevant Human Research Ethics Committee at RMIT University and all children and their parents gave their informed consent. Children were not provided an incentive for participation. All children were tested in schools, during school hours, in a quiet room. Administration of testing sessions (completed as part of a separate longitudinal study, see Ruddock et al. (2016)) was conducted over two 45-min sessions, with tasks administered in a counterbalanced order. In addition to the demographic questionnaire and motor screening assessment, participants completed the Double-Step Reaching Task (Hyde & Wilson, 2013) and tests of visual processing speed (Cogstate, 2018); however, these additional tasks were not included in the current analysis. For the GMLT, each child sat in a chair at a desk while the tablet PC was centred in front of them, with the screen positioned vertically and level with head height, close enough for their arm to comfortably reach and touch the screen with a stylus pen. Children were told the rules of the GMLT and then provided with untimed practice trials using a novel pathway. Additional practice trials were provided, if necessary, until it was determined that the child understood the task.

Design and data analysis

To examine age-related changes on GMLT metrics over childhood, we used a *cohort sequential design* (CSD) and estimated developmental trends over the 5.5-to-12.5-year-old period using growth curve modeling (GCM). A CSD combines a sequence of separate age cohorts to form a single overlapping age distribution, and is used to generate longitudinal data when performance is measured over a limited temporal scale (Duncan et al., 1996; Estrada & Ferrer, 2019; Prinzie & Onghena, 2005). In the current study, children were classified into one of 12 age cohorts based on their age (in months) at Time 0, rounded to the nearest 6-month increment. Age cohorts were defined in 6-month increments, resulting in the following cohorts: 5.5-, 6-, 6.5-, 7-, 7.5-, 8-, 8.5-, 9-, 9.5-, 10-, 10.5-, and 11-years. Children were assessed at five time points, separated by 6 months, over a 2-year period. Older children in grades 5 and 6 were tested until they graduated from primary school. Growth curves were estimated to describe the pattern of age-related change on GMLT performance metrics over the 5.5- to 12.5-year age period. In our study, less than 10% of children failed to complete all five occasions of testing. Importantly, GCM is used to model nonlinear developmental trajectories where the rate of growth often varies as a function of age (Curran et al., 2010; DeLucia & Pitts, 2006).

Analytic approach

Outliers (values >±3 SDs from the mean score) were assessed at each of the five time points. Only two data points (<1% of the data) were determined to be outliers, and thus no outliers were excluded. This ensured that all valid data points contributed to the integrity of the model fit (Ruddock et al., 2016). To ensure that data sufficiently represented the sample tested in the polynomial GCM, each age cohort at the five time points contained observations from at least three different children (Curran et al., 2010). Statistical models were estimated using the *lme4* function in the *lmerTest* package in RStudio (Bates et al., 2015). Models were estimated for each outcome variable separately. Firstly, a null model (Model 0) including only the outcome variable and random intercept of participant ID was estimated. Terms were then added to models and tested hierarchically according to the following sequence: Model 1 random intercept of participant ID nested within cohort was added due to the hierarchical nature of the study design; Model 2 fixed effect of participant age was added; Model 3 fixed effect of testing period and interaction between participant age and testing period was added; Model 4 fixed effect of orthogonal (using the *poly* function) quadratic term for participant age was added; Model 5 fixed effect of orthogonal cubic term for participant age was added. Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to compare goodness of fit between models. The fit of each model was compared to the previous model (e.g., Model 1 vs. Model 2, Model 2 vs. Model 3, etc.) using the *anova* function, with a significant difference (i.e., $p < .05$) used to determine if model fit was improved. The most complex model, Model 5, is represented by:

$$\text{value}_i \sim N\left(\mu, \sigma^2\right),$$

$$\mu = \alpha_{j[i],k[i]} + \beta_1 \bullet \text{Age} + \beta_2 \bullet \text{Age}^2 + \beta_3 \bullet \text{Age}^3 + \gamma_k \bullet I(\text{Period} = k, 1 \ldots 5) + \beta_j \bullet \left(\text{Age} \bullet \gamma_k\right),$$

where $\alpha_j \sim N\left(\mu_{\alpha j}, \sigma^2_{aj}\right)$ for ID: Cohort; where $\alpha_k \sim N\left(\mu_{\alpha k}, \sigma^2_{ak}\right)$ for Cohort.

Value was the outcome variable, with age and testing period used as fixed effects as well as their interaction. Orthogonal polynomial quadratic and cubic terms were included as additional fixed effects. A nested random effects structure was specified with id clustered within each cohort. Detailed models and their output are provided in Appendix S1.

Total errors on the GMLT for children aged 6–9 years at Time 0 was compared with normative data provided by Cogstate (2018) using independent t-tests.

RESULTS

A detailed breakdown of the number of observations, mean, SD, SE, and confidence intervals according to participant age, pooled across cohorts, for each GMLT variable is provided in Appendix S2. Results for response duration are presented in Appendix S3.

Legal errors

Participant legal errors, colored by cohort, and model estimate with 90% prediction interval, are presented in Figure 2.

Descriptive analysis

Visual inspection suggests that legal errors decline at a steady rate over the course of childhood; however, there was some slight between-age fluctuation in performance from 8.5 to 12.5 years.

Growth curve analysis

A quadratic trend was found to be the best fitting model for GMTL legal errors, $p = .03$; $-2\text{LL} = 2155.7$, AIC = 4339, BIC = 4402. Adding a cubic term did not improve model fit significantly, $p = .31$, therefore, a quadratic curve was deemed to have the most parsimonious fit (Table 2).

Rule-break errors

Participant rule-break errors, colored by cohort, and model estimate with 90% prediction interval, are presented in Figure 3.

Descriptive analysis

Visual inspection reveals a steep reduction in the number of rule-break errors between 5.5 and 7 years, followed by a more gradual reduction until 9 years, followed by a flattening of the curve up to 12.5 years.

Growth curve analysis

A quadratic trend was found to be the best fitting model for GMTL rule-break errors, $p < .001$; $-2\text{LL} = 1900.8$, AIC = 3830, BIC = 3892. Adding a cubic term did not improve model fit significantly, $p = .75$; therefore, a quadratic curve was deemed to have the most parsimonious fit (Table 2).

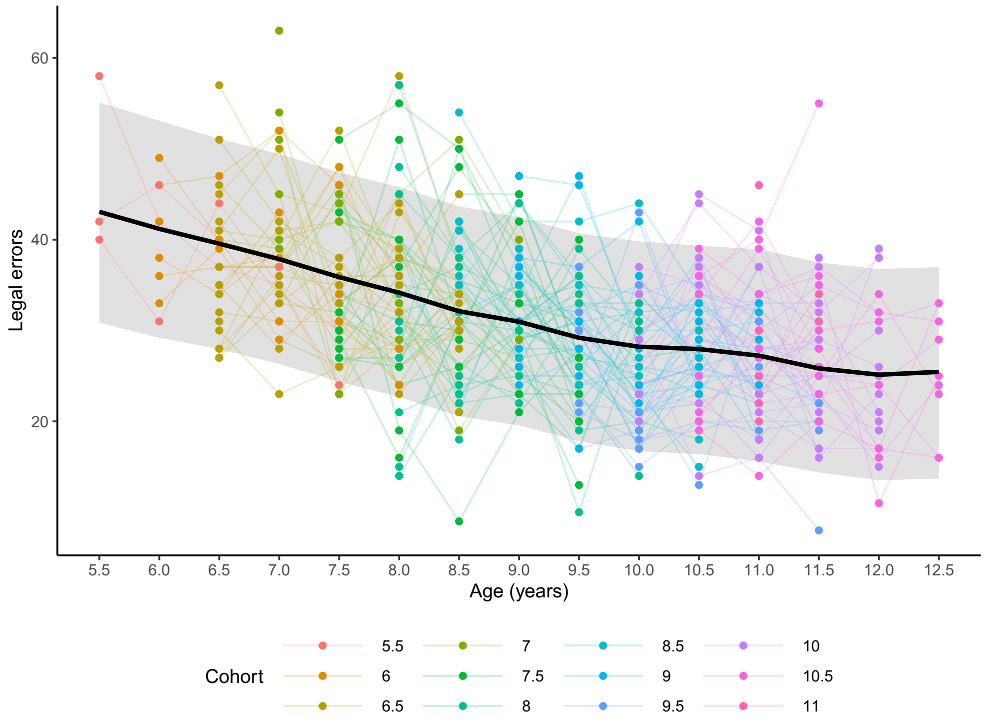

FIGURE 2 Participant legal errors, colored by cohort, made when completing the Groton Maze Learning Task. Black line indicates growth curve model estimate with 90% prediction interval.

Total errors

Participant total errors, colored by cohort, and model estimate with 90% prediction interval, are presented in Figure 4.

Descriptive analysis

Total errors on the GMLT at Time 0 did not differ between our sample and the standardization sample of Cogstate (Cogstate, 2018) at age 6, 7, 8, or 9 years (each $t<1$, $p>.05$). Visual inspection suggests that the average number of total errors made by children on the GMLT decreased at a steady rate over the 5.5 to 10-year age period, with the suggestion of more modest reduction after this time period.

Growth curve analysis

A quadratic trend was found to be the best fitting model for GMTL total errors, $p<.001$; $-2LL=2361.3$, $AIC=4751$, $BIC = 4813$. Adding a cubic term did not improve model fit significantly, $p = .54$; therefore, a quadratic curve was deemed to have the most parsimonious fit (Table 2).

DISCUSSION

The aim of this longitudinal study was to investigate and model the development of complex EF in healthy children using a rule-based, maze learning paradigm—the GMLT. Growth curve models were generated on each GMLT metric, and the fit compared between linear, quadratic, and cubic age using BIC and AIC criteria. On legal errors, rule-break errors, and total errors, quadratic trends were the best fitting and most parsimonious model in each case. For changes in response duration over childhood, a cubic trend was the best fitting (and more parsimonious) growth function. As predicted, we showed a quadratic growth trend on error monitoring (i.e., rule-break errors). Contrary to predictions, a quadratic was also shown on a metric of visuospatial working memory (i.e., legal errors), however the shape of this function was much shallower over the age period studied (5.5–12.5 years). The pattern of steep decline in rule-break errors between 5.5 and 7 years suggests a period of rapid consolidation of simple action-oriented rules that support learning of a visuospatial navigation task. These results are discussed below in relation to other paradigms commonly used to assess complex EF in children. The implications of our findings for developmental theory are also considered.

The number of total errors committed by children in our cohort did not differ from the standardized sample reported by Cogstate (2018), indicating that GMLT performance in our sample was representative of typically developing children. A rapid reduction in total response duration was found between 6.5 and 8.5 years followed by further (but more gradual) reduction until 12.5 years. This change in response duration reflects an

TABLE 2 Growth curve model fit performance for Groton Maze Learning Task variables

Age trend	Response duration				Total errors				Legal errors				Rule-break errors			
	AIC	BIC	−2LL	*p*	AIC	BIC	−2LL	*p*	AIC	BIC	−2LL	*p*	AIC	BIC	−2LL	*p*
Linear[a]	6245.15	6303.05	3109.58	<.01	4758.43	4816.33	2366.22	<.01	4341.99	4399.89	2158	.06	3842.31	3900.21	1908.16	<.01
Quadratic[b]	6237.31	6299.66	3104.66	<.01	**4750.61**	**4812.96**	**2361.3**	**<.01**	**4339.45**	**4401.8**	**2155.73**	**.03**	**3829.69**	**3892.04**	**1900.84**	**<.01**
Cubic[c]	**6233.46**	**6300.26**	**3101.73**	**.02**	4752.24	4819.04	2361.12	.54	4340.42	4407.23	2155.21	.31	3831.59	3898.39	1900.79	.75

Note: Bold font indicates model deemed to have most parsimonious fit.

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion.

[a]Comparison between Model 2 and Model 3.

[b]Comparison between Model 3 and Model 4.

[c]Comparison between Model 4 and Model 5.

improvement in performance efficiency and psychomotor speed, in line with (1) the reduction in rule-break and legal error rates, interpreted as a reduced time–cost in the integration and application of rules, and (2) improvements in fine- and gross-motor capabilities, which facilitate precise and rapid motor actions required to physically interact with the maze quickly.

The shallow quadratic trend on our measure of visuospatial working memory (or legal/spatial errors) is consistent with age trends reported for working memory tasks like the Corsi blocks (Farrell Pagulayan et al., 2006) and backward color recall (Röthlisberger et al., 2013). Indeed, quadratic trends have been suggested for more complex working memory tasks like *backward color recall* (Röthlisberger et al., 2013), which requires the recital of a series of colors in reverse order. Taken together, quadratic trends in the development of (visuospatial) working memory are likely to manifest under both simple and complex task instructions, including integration of a higher-order action-oriented instruction, such as reversing the order of presented stimuli. The current study provides considerable strength in its repeated use of the same task and longitudinal modeling of working memory over childhood; however, studies that sample over a wider age range (i.e., beyond 12 years old) are required to better understand developmental trends on more complex tasks.

The quadratic trend that we observed on rule-break errors suggests that the ability to use simple action-oriented rules in a spatial learning task is acquired rapidly over the course of early-to-middle childhood. Children reduced the average number of such errors from 20 at 5.5 years of age, to less than 10 errors by age 7, and to around four errors by 9 years. The ability to apply such rules is thought to be supported by a process of self-reflection and higher-order attentional control (Zelazo, 2015). Even use of simple rules, such as those involved in the GMLT, requires a capacity to monitor goal-directed behavior in real-time. In essence, reflection (via sub-vocal verbalisation) is thought to enable the child to reference previously learned task rules in order to solve an action problem, while also keeping other features of the task in mind (e.g., accessing the correct path held in working memory; Zelazo, 2015). An inability to reflect may also explain the utilization deficiency that occurs with young children, where a gap exists between the ability to learn task rules and call upon them mid-task (Clerc et al., 2014). The improvement with age may be further aided by the transition from reactive to proactive cognitive control that occurs from 3- to 8-years of age (Munakata et al., 2012). In the context of the GMLT, it may be that older children are better able to proactively maintain action-oriented rules for future planning of moves, whereas when this cognitive control is enlisted in younger children, rule use is slower and more cognitively demanding, resulting in more rule-break errors and slower overall response time. Taken together,

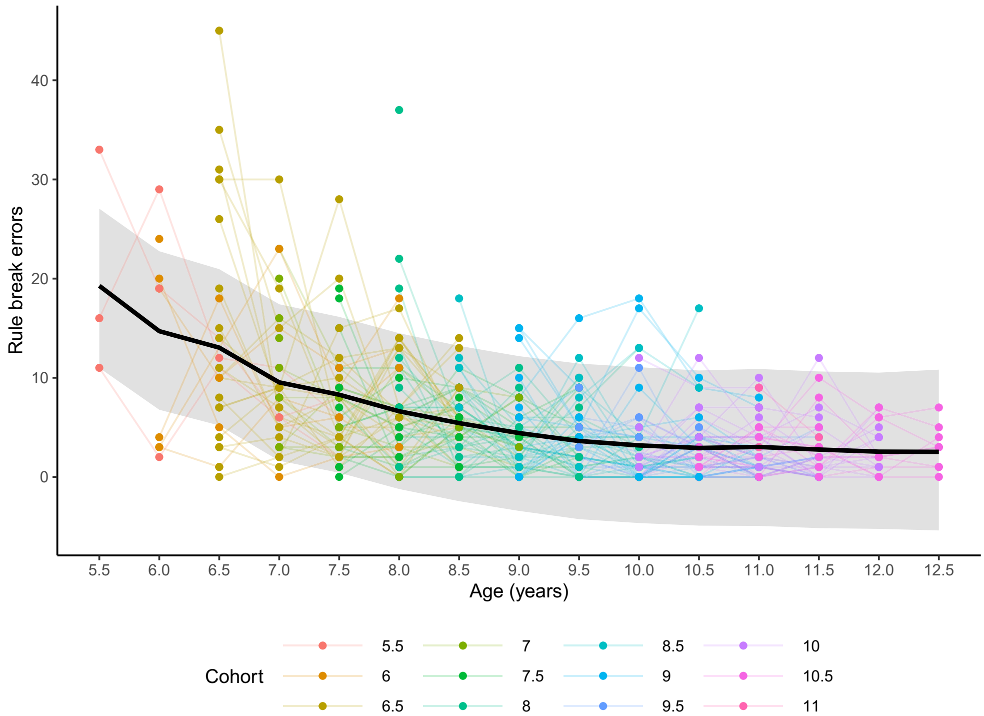

FIGURE 3 Participant rule-break errors, colored by cohort, made when completing the Groton Maze Learning Task. Black line indicates growth curve model estimate with 90% prediction interval.

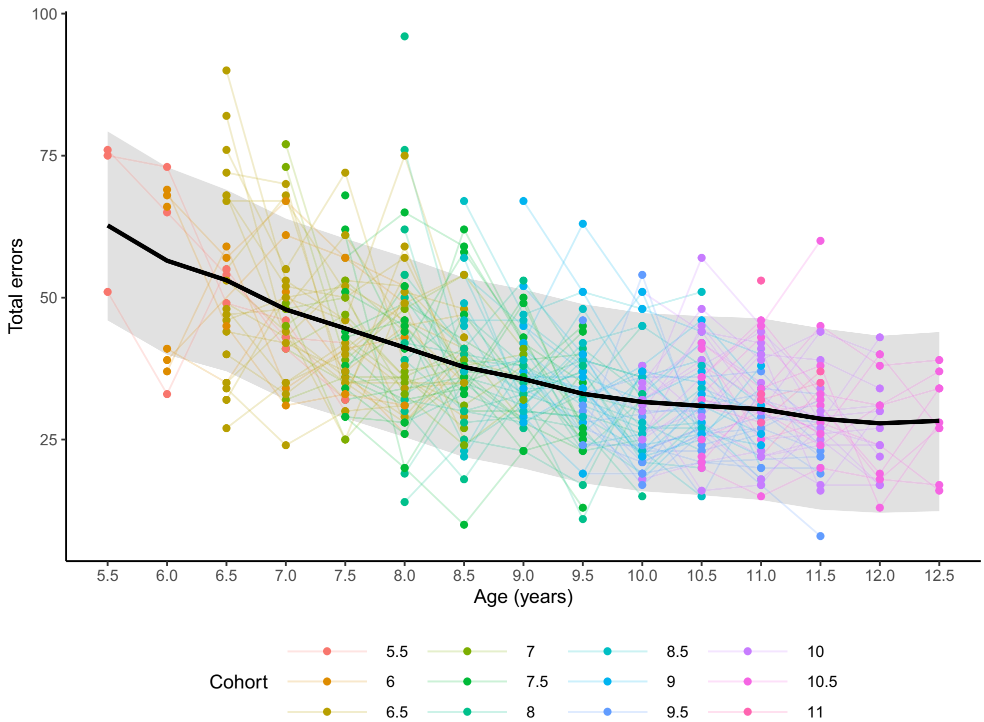

FIGURE 4 Participant total errors, colored by cohort, made when completing the Groton Maze Learning Task. Black line indicates growth curve model estimate with 90% prediction interval.

we argue that age-related changes in the ability to enlist simple action-oriented rules through reflection and proactive control, combined with an expanding capacity to buffer visuospatial information in working memory, maintains steady improvement on the GMLT over this critical developmental period.

It follows that performance of visuospatial learning tasks that involve more complex rule structures will follow different growth trajectories to those with simpler rule structures, like the GMLT (Perone et al., 2018). This is supported by both structural and functional neuroimaging data. First, tasks that involve hierarchical rule structures or more abstract reasoning tend to rely more heavily on lateral regions of the PFC (Bunge & Zelazo, 2006). These higher-order cortical regions develop relatively late, and then exert top-down influence on lower-level (but earlier developing) sensorimotor regions. Second, at a neural level, there is a shift with age and experience from sparse (within network) activation to a more targeted and well-defined grouping of network activation (Chevalier et al., 2019). This changing pattern occurs progressively over childhood and into adolescence. Finally, structural immaturities in the development of EF networks in younger children appear to necessitate a different pattern of neural activation. Put simply, when presented with a complex spatial planning and updating task, younger children appear to enlist networks other than those associated with mature EF—the left parietal cortex is one region where compensatory activation is evident (Morton et al., 2009). Taken together, our behavioral modeling of EF abilities suggests that performance of younger children may be supported by these less-efficient compensatory processes that are yet to fully develop. However, higher-level control is rapidly conferred as children enter middle-to-late childhood. With age, the capacity to couple the rule-monitoring functions of the PFC (and cingulate cortex) with lower-level spatial buffering in working memory is likely to underpin advances in complex task performance.

Limitations

Despite the strengths of the current investigation, there are some limitations that should be considered when interpreting the findings. Primarily, considering that children with MAND scores below the 20th percentile were excluded, the current findings cannot be extended to children with developmental delays in motor skill. This is an important consideration given the EF deficits (e.g., inhibitory control) of children with motor difficulties, such as Developmental Coordination Disorder (Subara-Zukic et al., 2022; Wilson et al., 2013, 2017). Furthermore, given the lack of more detailed demographic information, the generalisability of the findings are somewhat limited to typically developing children in metropolitan areas of Western-cultured cities of Melbourne and Perth, Australia. It is likely that developmental trends in EF will be similar for typically developing cohorts arising from other metropolitan (and regional) areas; however, without the availability of more comprehensive demographic data, this cannot be assumed. Finally, as the GMLT was administered as part of a battery of tests in a separate longitudinal study (Ruddock et al., 2016), it is possible that fatigue and/or motivation for the task could have varied between timepoints. However, as the order of assessments were counterbalanced across time points, this should mitigate the issue of learning effects between 6-month testing points, and the results may actually give a more representative insight into day-to-day EF of children.

Future directions

One factor for future investigation concerns possible cross-cultural differences in the development of EF between developing and emerging countries. There is some evidence that growth patterns in EF are not as readily observed in developing regions (Schirmbeck et al., 2020). Cross-cultural comparisons are needed to fully understand growth patterns and the constraints of living conditions, socio-economic status, diet, and physical activity.

Looking beyond pre-set developmental trajectories (i.e., linear, quadratic and cubic), future modeling of growth curves could be extended by testing the best fitting function without regard to its shape (termed spline function) which help model critical points along a developmental trajectory where change occurs. These functions maximize data fit but are often difficult to interpret in terms of current theories of development (Telzer et al., 2018).

CONCLUSION

Typically developing children show distinct patterns of growth in the ability to enlist different facets of EF in a rule-based learning task. The decline on rule-break errors between 5.5 and 9 years combined with quadratic growth trend over the childhood period suggests rapid acquisition of the ability to enlist simple action-oriented rules in goal-directed behavior (specifically, maze navigation). This change in cognitive control is likely to act as a rate limiting factor in performance of tasks that require complex EF. These behavioral results mirror structural changes in the development of neural networks that underpin working memory and cognitive control, particularly synergies between PFC and cingulate cortex (Botvinick & Cohen, 2014; Garner & Dux, 2020). Steady maturation in such networks supports the ability to integrate the maintenance of visual–spatial information online while solving cognitive tasks. For tasks that involve more complex action-oriented rules (than those examined here), a more protracted period of development may be necessary (Zelazo, 2015). The ability to describe these patterns of growth in EF with greater precision has important implications for the design of learning environments for school-aged children; the aim of which should be to present learning tasks that impose constraints on working memory and cognitive control that are scaled appropriately for proximal learning.

ACKNOWLEDGMENTS

We thank the Australian Research Council (ARC) for funding this project under the Discovery Grants Scheme (DP1094535). The data and analytic code necessary to reproduce the analyses presented here are publicly accessible. Data and analytic code are available from the first author upon reasonable request. The materials necessary to attempt to replicate the findings presented here are not publicly accessible. The analyses presented here were not preregistered. Open access publishing facilitated by Australian Catholic University, as part of the Wiley - Australian Catholic University agreement via the Council of Australian University Librarians.

ORCID

Thomas B. McGuckian https://orcid.org/0000-0002-5490-0042
Peter H. Wilson https://orcid.org/0000-0003-3747-0287
Rich D. Johnston https://orcid.org/0000-0001-6618-2853
Shahin Rahimi-Golkhandan https://orcid.org/0000-0001-7566-2445
Jan Piek https://orcid.org/0000-0003-3838-6773
Dido Green https://orcid.org/0000-0002-1129-8071
Jeffrey M. Rogers https://orcid.org/0000-0002-0320-969X
Paul Maruff https://orcid.org/0000-0002-6947-9537
Bert Steenbergen https://orcid.org/0000-0001-8863-2624
Scott Ruddock https://orcid.org/0000-0002-8245-1205

REFERENCES

Alloway, T. P., & Gathercole, S. E. (2012). *Working memory and neurodevelopmental disorders*. https://doi.org/10.4324/9780203013403

Australian Bureau of Statistics. (2021). *Census of population and housing: Cultural diversity data summary, 2021*. Australian Government.

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using LME4. *Journal of Statistical Software*, *67*(1), 48–96. https://doi.org/10.18637/jss.v067.i01

Benson, J. E., Sabbagh, M. A., Carlson, S. M., & Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Developmental Psychology*, *49*(9), 1615–1627. https://doi.org/10.1037/a0031056

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336. https://doi.org/10.1016/j.lindif.2011.01.007

Botvinick, M. M., & Cohen, J. D. (2014). The computational and neural basis of cognitive control: Charted territory and new frontiers. *Cognitive Science*, *38*(6), 1249–1285. https://doi.org/10.1111/cogs.12126

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, *15*(3), 118–121. https://doi.org/10.1111/j.0963-7214.2006.00419.x

Carlson, S. M., Zelazo, P. D., & Faja, S. (2019). From executive function to executive functions: A neurocognitive and socio-affective synthesis in cognitive neurosciences and developmental psychology. *ANAE—Approche Neuropsychologique des Apprentissages chez l'Enfant*, *31*(160), 371–413. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85069505531&partnerID=40&md5=1466543eb4d9f73ed41a738294ea6c39

Casey, B. J., Galvan, A., & Hare, T. A. (2005). Changes in cerebral functional organization during cognitive development. *Current Opinion in Neurobiology*, *15*, 239–244.

Chevalier, N., Jackson, J., Revueltas Roux, A., Moriguchi, Y., & Auyeung, B. (2019). Differentiation in prefrontal cortex recruitment during childhood: Evidence from cognitive control demands and social contexts. *Developmental Cognitive Neuroscience*, *36*, 100629. https://doi.org/10.1016/j.dcn.2019.100629

Clerc, J., Miller, P. H., & Cosnefroy, L. (2014). Young children's transfer of strategies: Utilization deficiencies, executive function, and metacognition. *Developmental Review*, *34*, 16–393.

Cogstate. (2018). *Cogstate pediatric and adult normative data*. Cogstate Limited.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121–136. https://doi.org/10.1080/15248371003699969

DeLucia, C., & Pitts, S. C. (2006). Applications of individual growth curve modeling for pediatric psychology research. *Journal of Pediatric Psychology*, *31*(10), 1002–1023. https://doi.org/10.1093/jpepsy/jsj074

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. *Psychological Methods*, *1*(3), 236–248.

Estrada, E., & Ferrer, E. (2019). Studying developmental processes in accelerated cohort-sequential designs with discrete-and continuous-time latent change score models. *Psychological Methods*, *24*(6), 708–734.

Farrell Pagulayan, K., Busch, R., Medina, K., Bartok, J., & Krikorian, R. (2006). Developmental normative data for the Corsi block-tapping task. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 1043–1052. https://doi.org/10.1080/13803390500350977

Garner, K. G., & Dux, P. E. D. (2020). The neural basis of multitasking. In K. G. G. D. Lloyd (Ed.), *Visualisations*. https://doi.org/10.31234/osf.io/jqxeb

Hyde, C., & Wilson, P. H. (2013). Impaired online control in children with developmental coordination disorder reflects developmental immaturity. *Developmental Neuropsychology*, *38*(2), 81–97. https://doi.org/10.1080/87565641.2012.718820

McCarron, L. T. (1997). *McCarron Assessment of Neuromuscular Development* (Rev. ed.). Common Market Press.

Mirabella, G. (2021). Inhibitory control and impulsive responses in neurodevelopmental disorders. *Developmental Medicine and Child Neurology*, *63*(5), 520–526. https://doi.org/10.1111/dmcn.14778

Morton, J. B., Bosma, R., & Ansari, D. (2009). Age-related changes in brain activation associated with dimensional shifts of attention: An fMRI study. *NeuroImage*, *46*(1), 249–256. https://doi.org/10.1016/j.neuroimage.2009.01.037

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science*, *21*, 7–77.

Pascual, A. C., Moyano, N., & Robres, A. Q. (2019). The relationship between executive functions and academic performance in primary education: Review and meta-analysis. *Frontiers in Psychology*, *10*(July), 1582. https://doi.org/10.3389/fpsyg.2019.01582

Perone, S., Almy, B., & Zelazo, P. D. (2018). Toward an understanding of the neural basis of executive function development. In *The neurobiology of brain and behavioral development* (pp. 291–314). Elsevier Inc. https://doi.org/10.1016/B978-0-12-804036-2.00011-X

Pietrzak, R. H., Maruff, P., Mayes, L. C., Roman, S. A., Sosa, J. A., & Snyder, P. J. (2008). An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring. *Archives of Clinical Neuropsychology*, *23*(4), 433–445. https://doi.org/10.1016/j.acn.2008.03.002

Prinzie, P., & Onghena, P. (2005). Cohort sequential design. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (pp. 319–322). John Wiley & Sons.

Röthlisberger, M., Neuenschwander, R., Cimeli, P., & Roebers, C. M. (2013). Executive functions in 5- to 8-year olds: Developmental changes and relationship to academic achievement. *Journal of Educational and Developmental Psychology*, *3*, 16.

Ruddock, S., Caeyenberghs, K., Piek, J., Sugden, D., Hyde, C., Morris, S., Rigoli, D., Steenbergen, B., & Wilson, P. (2016). Coupling of online control and inhibitory systems in children with atypical motor development: A growth curve modelling study. *Brain and Cognition*, *109*, 84–95. https://doi.org/10.1016/j.bandc.2016.08.001

Schirmbeck, K., Rao, N., & Maehler, C. (2020). Similarities and differences across countries in the development of executive functions in children: A systematic review. *Infant and Child Development*, *29*(1), e2164. https://doi.org/10.1002/icd.2164

Snyder, A. M., Maruff, P., Pietrzak, R. H., Cromer, J. R., & Snyder, P. J. (2008). Effect of treatment with stimulant medication on nonverbal executive function and Visuomotor speed in children with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *14*, 16.

Snyder, P. J., Jackson, C. E., Piskulic, D., Olver, J., Norman, T., & Maruff, P. (2008). Spatial working memory and problem solving in schizophrenia: The effect of symptom stabilization with atypical antipsychotic medication. *Psychiatry Research*, *160*, 11.

Snyder, P. J., Werth, J., Giordani, B., Caveney, A. F., Feltner, D., & Maruff, P. (2005). A method for determining the magnitude of change across different cognitive functions in clinical trials: The effects of acute administration of two different doses alprazolam. *Human Psychopharmacology*, *20*(4), 263–273. https://doi.org/10.1002/hup.692

Subara-Zukic, E., Cole, M., McGuckian, T., Steenbergen, B., Green, D., Smits-Engelsman, B., Lust, J. M., Abdollahipour, R., Domellof, E., Deconinck, F., Blank, R., & Wilson, P. H. (2022). Behavioural and neuroimaging research on developmental coordination disorder (DCD): A combined systematic review and meta-analysis on recent findings. *Frontiers in Psychology*, *13*, 1–28.

Telzer, E. H., McCormick, E. M., Peters, S., Cosme, D., Pfeifer, J. H., & van Duijvenvoorde, A. C. K. (2018). Methodological considerations for developmental longitudinal fMRI research. *Developmental Cognitive Neuroscience*, *33*, 149–160. https://doi.org/10.1016/j.dcn.2018.02.004

Thomas, E., Maruff, P., Paul, J., & Reeve, R. (2016). Spatial sequence memory and spatial error monitoring in the Groton Maze Learning Task (GMLT): A validation study of GMLT submeasures in healthy children. *Child Neuropsychology*, *22*(7), 837–852. https://doi.org/10.1080/09297049.2015.1038989

Thomas, E., Reeve, R., Fredrickson, A., & Maruff, P. (2011). Spatial memory and executive functions in children. *Child Neuropsychology*, *17*(6), 599–615. https://doi.org/10.1080/09297049.2011.567980

Thomas, E., Reeve, R., Pietrzak, R., & Maruff, P. (2013). Disentangling component learning and executive processes in hidden pathway maze learning in children: A process-based approach. *Child Neuropsychology*, *19*(6), 588–600. https://doi.org/10.1080/09297049.2012.704010

Thomas, E., Snyder, P. J., Pietrzak, R. H., & Maruff, P. (2014). Behavior at the choice point: Decision making in hidden pathway maze learning. *Neuropsychology Review*, *24*(4), 514–536. https://doi.org/10.1007/s11065-014-9272-7

Tirapu-Ustarroz, J., Bausela Herreras, E., & Cordero-Andrés, P. (2018). Model of executive functions based on factorial analyses in child and school populations: A meta-analysis. *Revista de Neurologia*, *67*, 215–225.

Wilson, P. H., Ruddock, S., Smits-Engelsman, B., Polatajko, H., & Blank, R. (2013). Understanding performance deficits in developmental coordination disorder: A meta-analysis of recent research. *Developmental Medicine and Child Neurology*, *55*(3), 217–228. https://doi.org/10.1111/j.1469-8749.2012.04436.x

Wilson, P. H., Smits-Engelsman, B., Caeyenberghs, K., Steenbergen, B., Sugden, D., Clark, J., Mumford, N., & Blank, R. (2017). Cognitive and neuroimaging findings in developmental coordination disorder: New insights from a systematic review of recent research. *Developmental Medicine and Child Neurology*, *59*(11), 1117–1129. https://doi.org/10.1111/dmcn.13530

Zelazo, P. D. (2015). Executive function: Reflection, iterative reprocessing, complexity, and the developing brain. *Developmental Review*, *38*, 55–68. https://doi.org/10.1016/j.dr.2015.07.001

Zelazo, P. D., & Carlson, S. M. (2017). Embracing complexity in the study of executive function and its development. In *Advancing developmental science: Philosophy, theory, and method* (pp. 110–118). Taylor and Francis. https://doi.org/10.4324/9791315174686

Zelazo, P. D., & Carlson, S. M. (2020). The neurodevelopment of executive function skills: Implications for academic achievement gaps. *Psychology & Neuroscience*, *13*(3), 26. https://doi.org/10.1037/pne0000208

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: McGuckian, T. B., Wilson, P. H., Johnston, R. D., Rahimi-Golkhandan, S., Piek, J., Green, D., Rogers, J. M., Maruff, P., Steenbergen, B., & Ruddock, S. (2023). Development of complex executive function over childhood: Longitudinal growth curve modeling of performance on the Groton Maze Learning Task. *Child Development*, *94*, 648–658. https://doi.org/10.1111/cdev.13888