Child Development, 2026, 1–17
https://doi.org/10.1093/chidev/aacag002
Published: February 26, 2026
Empirical Article

Tracking the trajectory of executive function from 2.5 to 6.5 years of age and the impact of COVID-19

Eleanor Johns[1,2], Samuel H. Forbes[3], Lourdes M. Delgado Reyes[2], Charlotte Buck[2], and John P. Spencer[2,*]

[1]School of Psychology, Lancaster University, Lancaster, United Kingdom
[2]Schoool of Psychology, University of East Anglia, Norwich, United Kingdom
[3]School of Psychology, Durham University, Durham, United Kingdom
*Corresponding author: School of Psychology, University of East Anglia, Norwich NR4 7TJ, United Kingdom. Email: j.spencer@uea.ac.uk (J.P.S.)

Abstract

Given the significance of executive function (EF) for longer-term outcomes, it is important to understand the trajectory of EF in childhood and how COVID-19 influenced this trajectory. 139 (71 girls) children were examined longitudinally from 2.5 to 6.5 years using the Minnesota Executive Function Scale. Individual differences in EF abilities were stable longitudinally, and children with a more highly educated mother had higher EF scores. Children tested 1–3 years after the first lockdown showed greater variation in individual differences over time. Additional analyses examined linear growth curves for children assessed 3–4 times from 2.5 to 6.5 years. Children who were in preschool during the first lockdown had steeper growth curve trajectories than children in the first year of primary school.

Keywords executive function, COVID-19, cognition, developmental trajectories

Lay summary

We studied children's self-regulation skills across early childhood focusing on executive functions. This was disrupted by COVID-19, prompting us to investigate COVID-19's influence on children's executive functions. One hundred and thirty-nine (71 girls, 68 boys) children visited our lab multiple times from 2.5 to 6.5 years of age, with 94 families first attending before COVID-19. Using the Minnesota Executive Function Scale, we found children's executive functioning increased consistently from 2.5 to 6.5 years of age but was affected by COVID-19. We found children who had already started primary school showed slower gains in executive functioning relative to younger children who were in preschools at the start of the pandemic, likely reflecting disruptions to early schooling and social routines during this critical adjustment period.

Executive function (EF) is a central cognitive system that enables self-control and flexibility, helping children to give a considered rather than impulsive response and to stay on-task (Diamond, 2013). Although there are differences in how EF is conceptualized (Howard et al., 2015; Miyake et al., 2000; Wiebe et al., 2011), one common view posits that EF consists of a set of higher-level cognition processes that sit above and coordinate lower-level component skills that include working memory, inhibition control, and cognitive flexibility (Miyake et al., 2000). As an example, some EF tasks ask people to follow one set of rules and then to switch and follow a different set of rules. Executive function would keep the higher-level rules of the task active in each phase and switch from one rule set to the other as the situation changes.

Over the last two decades, EF has become an increasingly important topic as evidence suggests EF is strongly linked to academic achievement and longer-term outcomes (Carlson, 2021; Carlson et al., 2004; Moffitt et al., 2011). For instance, EFs predict math and reading competence across the school years (Ahmed et al., 2019; Bull et al., 2008; Gathercole et al., 2004), as well as career and marriage satisfaction and positive mental and physical health in adulthood (Dunn, 2010; Eakin et al., 2004; Moffitt et al., 2011; Prince et al., 2007; Richmond-Rakerd et al., 2021). Given this long-term predictability, a key issue is how EFs change in early development. If we could optimize EF early in development—ensuring that every child has good EF skills by 3–5 years—this could have a positive societal impact (Heckman et al., 2006).

To achieve robust early assessment and design effective interventions, we need tasks that are optimized for these goals. This also requires longitudinal assessment from early childhood, which has been relatively rare in the literature (see Broomell & Bell, 2022 and Willoughby et al., 2012 for exceptions). Finally, we need to understand the mechanisms that underlie the early

Handling Editor: Seth Pollak
Editor: Shauna Cooper
Received: August 6, 2025. **Revised:** December 28, 2025. **Accepted:** December 31, 2025

development of EF (Buss & Spencer, 2014; McCraw et al., 2024; Spencer et al., 2025). Meeting these goals has proven to be difficult for several reasons.

First, children are rapidly changing in early development. Consequently, EF tasks that are easy enough early in development become too easy later in development. For instance, during the first year of life, infants begin to search for hidden objects and make the canonical A-not-B error (Piaget, 1954; Thelen et al., 2001). Here, they perseverate by searching at an "A" location after an object that was repeatedly found in location "A" has now been hidden at a nearby "B" location (Bell, 2012; Diamond, 1990). Critically, individual differences in performance in this task are related to increased fronto-parietal activation from as young as 8 months (Bell, 2001), and longitudinal studies have shown that A-not-B performance in infancy is correlated with planning and executive control later (Wiebe et al., 2010), and predicts aspects of later EF when included in a task battery (Broomell & Bell, 2022). Although this is the case, the use of this task as an early assessment tool is limited by a lack of longitudinal stability in this measure. Miller and Marcovitch (2015) found no correlation in performance on the A-not-B task from 14 to 18 months of age. Moreover, while the A-not-B task is appropriate in early infancy, it reaches ceiling levels from around 24 months of age. After this age, other tasks are required to examine EF performance longitudinally (Broomell & Bell, 2022).

Several longitudinal studies have faced these obstacles by using different tasks at different ages. This presents a theoretical challenge as we do not always understand how different tasks relate to one another. As an example, the A-not-B task is often paired with response inhibition tasks such as Stroop-like and "go/no-go" tasks in longitudinal studies (Broomell & Bell, 2022; Johansson et al., 2012). Although the A-not-B task does involve response inhibition (i.e., infants have to inhibit responding to A after a cue to B), theoretical models have suggested that working memory is also involved as infants must actively maintain a memory of the "B" hiding event on the B trials (Thelen et al., 2001). This differs from the conceptualization of tasks such as the "go/no-go" task which tend to focus solely on inhibitory processes (although see Wijeakumar et al., 2017 for a discussion). Moreover, Stroop-like and "go/no-go" tasks involve language, while the A-not-B task is nonverbal. These differences make it difficult to interpret results suggesting that performance on the A-not-B task in infancy is related to later performance on a Stroop-like task in the toddler period (Johansson et al., 2012), as we do not know precisely what is changing over time to produce this relationship.

The theoretical picture gets even more complicated in early development when we consider evidence that EF and its lower-level components are co-developing over time. For instance, we see periods of vast improvement in young children's working memory, cognitive flexibility, and inhibitory control abilities during early childhood (Diamond, 2013; Fuhs et al., 2014; Garon et al., 2008). This makes it difficult to know if changes in EF are driven by improvements in, say, working memory or inhibitory control vs. changes in higher-level EF processes themselves such as in how the higher-level rules of a task are represented (for discussion, see Spencer et al., 2025).

One way forward empirically may be to use more complex tasks that tap into multiple EF components at once. Once language comprehension develops, toddlers can begin to play games in which they have to understand and keep multiple rules in mind and engage cognitive flexibility. This allows us to examine 2- to 5-year-old children's ability to switch between tasks (Brooks et al., 2003; Perner & Lang, 2002), inhibit actions (Steelandt et al., 2012) and maintain and use information held in working memory (Cheng et al., 2020; Garon et al., 2014). For instance, in the Dimensional Change Card Sort (DCCS) task (Frye et al., 1995), children sort cards by one dimension (e.g., shape) and then switch and sort cards by a second dimension (e.g., color). Theoretical models have shown that this task integrates aspects of cognitive flexibility, working memory, and inhibitory control (Buss & Spencer, 2014), and more recent work shows that aspects of language development—in particular, children's understanding of dimensional labels like "color" and "shape"—also contribute (Buss & Spencer, 2014; Lowery et al., 2022; Spencer et al., 2025). The link to language development here is interesting as this starts to unpack the higher-level cognitive processes at play in this particular EF task.

Evidence from studies using the DCCS task shows rapid change in performance from 3 to 5 years of age. Three-year olds are typically successful at sorting cards during a preswitch phase, but they continue to sort the test cards by the first dimension during a postswitch phase, although adaptations of this task have found 3-year olds are able to demonstrate successful switching when dimensions are more separated (Diamond et al., 2005) or the spatial location of the cards is changed postswitch (Buss & Spencer, 2014; Perone et al., 2015). By 5 years of age, children succeed in the postswitch phase without adaptation, although there is still improvement in performance beyond this age (Zelazo et al., 2013). For instance, the DCCS task has been developed into a tablet task—the Minnesota Executive Function Scale (MEFS; Carlson, 2021). In a recent cross-sectional examination of over 51,000 participants from different studies in the United States, performance on MEFS improved rapidly across early childhood, with more gradual improvements through adolescence (Carlson, 2021). Two longitudinal studies have used the DCCS task. The first examined EF from 18 to 26 months of age. At 18 months of age, Bernier et al. (2010) used a hide-the-pots task (an A-not-B style search task) and a DCCS style categorization task. These were compared with multiple tasks at 26 months of age including a delay-of-gratification task and two Stroop-like tasks. The hide-the-pots and DCCS style categorization tasks were found to have a low correlation at 18 months of age. Despite this, both tasks were found to be positively related to a measure of "conflict EF," which combined performance across multiple tasks at 26 months of age. Once again, it is difficult to make strong conclusions about the longitudinal stability of individual differences in EF from these results because this study used different tasks at different ages that may tap into different components of EF.

A second longitudinal study, by contrast, used the DCCS task consistently across ages with older children. Helm et al. (2020) found that children's performance on the DCCS at 4 years of age was positively correlated with performance on the DCCS at 6 years of age. Thus, at least with 4- to 6-year olds, individual differences in performance in this task are stable longitudinally.

The goal of the present study was to assess whether individual differences in EF abilities are stable over time before the age of 4 years using the DCCS task at each time point. We opted to use MEFS starting at 2.5 years of age (30 months)—the youngest age reliably tested using this assessment tool. We then tracked longitudinal performance using MEFS from 2.5 years (30 months) to 6.5

years (78 months). Our central question was whether we would observe robust individual differences across this time frame. We also examined whether these individual differences were influenced by two factors: socioeconomic status and the impact of the COVID-19 pandemic.

A large number of studies have examined how socioeconomic factors shape child development (Duncan et al., 2017; Hughes et al., 2009; Wiebe et al., 2011). A recent meta-analysis including 8,760 children found the correlation between EF and socioeconomic status was statistically significant across all studies (Lawson et al., 2018). Parental educational level has emerged as a strong predictor of children's cognitive outcomes from these studies. For instance, Waters et al. (2021) found that parental education level, and not income to needs, was associated with multiple EF and academic achievement domains, including math and reading ability. Thus, maternal education level was examined within the present study to examine whether this socioeconomic factor was associated with individual differences in EF abilities over time.

The second factor we examined was not preplanned—our longitudinal study was on-going when the SARS-CoV-2 (COVID-19) pandemic began. The pandemic brought widespread disruption to social, economic, and health environments, and it is likely EF development was also impacted. The government in the United Kingdom officially announced the first lockdown on March 23, 2020, where the public were asked to cease all nonessential travel and contact (Prime Minister's Office, 23 March 2020). Stress and social deprivation were commonplace during these lockdown restrictions, due to increased job losses and financial uncertainty. Evidence shows that such stressors are harmful to EF (Chichinina & Gavrilova, 2022). For instance, parent stress can impact parenting quality and emotional availability (Korzeniowski, 2023). As a result, parents may rely on control strategies rather than scaffolding techniques (Polizzi et al., 2021). Parents who experience frequent distress may engage less in cognitively challenging tasks or create a hostile environment, which can obstruct the development of a child's self-regulatory skills and overall cognitive growth (Polizzi et al., 2021).

Younger children were also impacted by closures to daycare centers, preschools, primary schools, and playgrounds. Exposure to such structured environments and socializing with peers supports EF development during the preschool and primary school years (Veraksa et al., 2024). When access to play and socialization is limited, EF abilities are adversely affected (Holmes et al., 2016). For instance, during the highly restrictive lockdown periods, children under three experienced significant reductions in cognitive abilities, scoring roughly 20 points lower on the Mullen Scales of Early Learning relative to children tested prior to the pandemic (Perry et al., 2023). Similarly, Navarro-Soria et al. (2023) reported a significant decline in global executive functioning just 6 months after the start of lockdown.

While various studies have explored the impact of the COVID-19 pandemic on EF (Chichinina & Gavrilova, 2022; Perry et al., 2023), the majority of research has been cross-sectional. Longitudinal research is necessary to track EF trajectories and determine whether early COVID-19 impacts recover, persist, or worsen as children age. Here, we provide a longitudinal examination of this issue in the cohort of children who were assessed at 30 months of age before the pandemic hit and again at 78 months of age. We examined two key questions. First, how were individual differences in EF impacted by COVID-19 based on the cumulative period from the onset of the first COVID-19 lockdown until the time when each child was assessed (which we refer to as the "months since COVID")? Here, we emphasize that all children were tested around 78 months of age; however, some children were tested 1–2 years after the first lockdown while other children were tested 3–4 years after the first lockdown. On one hand, we might expect that COVID-19 was most disruptive in the 1–2 years after the first lockdown as this period included multiple lockdowns and disruptions to social and educational routines. In this case, individual differences might be less stable longitudinally. Alternatively, we might expect that COVID-19 disrupted individual differences longer term. In this case, individual differences might be less robust the longer the time between the first COVID lockdown and testing at 78 months.

The second question we examined was how individual differences in EF trajectories were impacted by the age of each child at the onset of the pandemic (which we refer to as the "age at COVID onset"). Here, we examined EF trajectories over multiple time points (30, 42, 54, and 78 months) using data from a subset of children who completed three or more MEFS assessments. Some children were younger when the pandemic began (e.g., 45 months) while other children were older (e.g., 56 months). This might have important implications as the younger children were typically in preschool, while the older children were typically in the first year of primary school (i.e., reception). Preschool children might have been more profoundly impacted by the societal upheaval created by the pandemic as they might be more sensitive to early experiences. Alternatively, we might expect to see stronger impacts on primary school children because peer relationships—which were limited during the pandemic—play a key role as children move from preschool into primary school. In this case, we might expect to see slower growth of EF skills longitudinally in children who were older at the start of the pandemic.

Method

Study design

Data from the current study were collected as part of a larger longitudinal study examining the early development of working memory and EF. The larger study enrolled two cohorts of children. Cohort 1 was enrolled at 6 months of age. These children visited the lab for testing at 6, 18, and 30 months of age. Cohort 2 was enrolled at 30 months of age. These children visited the lab for testing at 30, 42, and 54 months of age. All children completed follow-up testing at 78 months of age.

The present report analyzed data from the MEFS task which were collected at different points in time for the two cohorts of children. Children in cohort 1 completed the MEFS task at 30 and 78 months. Children in cohort 2 completed the MEFS task at 30, 42, 54, and 78 months. Thus, when we combined the two cohorts, we had a large number of children who completed the MEFS task at 30 and 78 months. We also had a subset of children (cohort 2) who completed the MEFS task up to 4 times.

Because data from cohorts 1 and 2 were collected in parallel (i.e., we continuously enrolled children in cohorts 1 and 2 at the start of the study), this created natural variability in the age of each participant when the first COVID-19 lockdown was announced (i.e., age at COVID onset) as well as variation in the time elapsed between the first lockdown and the date of testing at 78 months

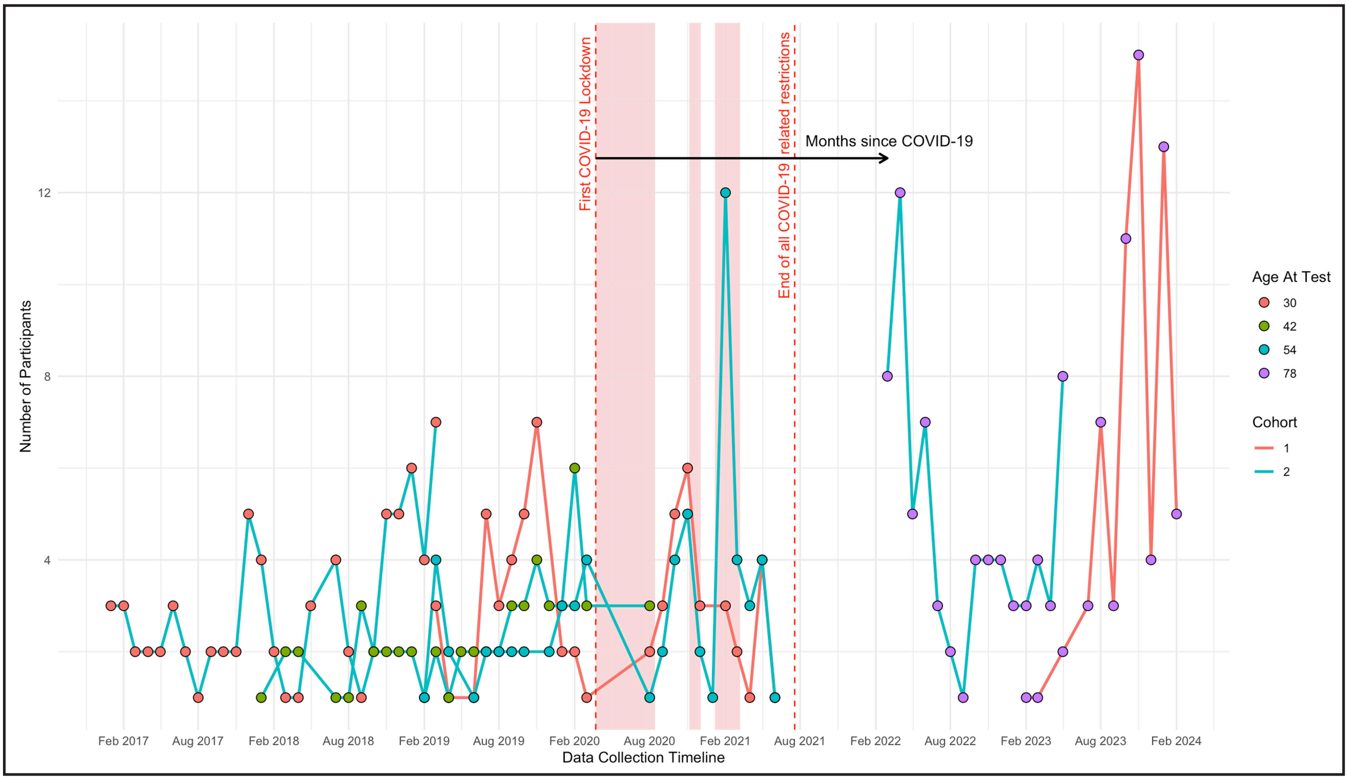

Figure 1 Timeline of data collection (see *x*-axis) with COVID-19 periods noted in red shading. Dot color shows the data collection ages (in months). Line color indicates cohort 1 (red; initially enrolled at 6 months of age) and cohort 2 (blue; initially enrolled at 30 months of age).

(i.e., months since COVID). Figure 1 shows the full timeline of MEFS data collection for the present report. The blue line shows MEFS data collection for cohort 1 which ran from February 2019 until June 2021, when the children were 30 months of age, and a second distribution from February 2023 until February 2024 when the children reached 78 months of age. The red line shows data collection for cohort 2. Here, there is a more continuous distribution of data collection indicated by the dot color.

Figure 1 also highlights the two key COVID-19-related measures we focus on in the analyses below. The first is the "age at COVID onset": this is the age of each child when the first COVID-19 lockdown began on March 23, 2020. As is evident in Figure 1, we had quite a variation in ages across cohorts 1 and 2 when the pandemic began. The second is a "months since COVID" measure: this is the duration in months from the first COVID-19 lockdown until data collection at 78 months. An example "months since COVID" interval is shown by the horizontal arrow in Figure 1.

Participants

One hundred and thirty-nine children (71 girls) from Norwich in the United Kingdom completed the EF tasks. Demographics for each cohort are shown in the first two columns of Table 1. Average maternal education level across cohorts was a Bachelor's Degree, and mean income was £40,645.69 (SD = 11,720.63). Participants had normal or corrected-to-normal vision. Color vision was examined through family history of color blindness risk; at-risk children were excluded. All participants were full-term infants with no history of early brain trauma. The project was reviewed and approved by the Ethics Committee at NHS England (IRAS ID 196063).

Parents signed an informed consent form on behalf of the child. Children received a toy and a t-shirt for participating at each lab visit. Parents were given £20 for each visit to the lab.

As described in Methods of analysis, we conducted five analyses in the present report—two analyses addressed whether individual differences in EF were stable longitudinally (analyses 1a and 2a) and three analyses examined how COVID-19 impacted these individual differences (analyses 1b, 2b, and 2c). Of the 139 children who participated in this study, 16 did not complete the MEFS task at 30 months of age and a further 3 children did not complete the follow-up MEFS task at 78 months of age. Thus, 120 children contributed data to analysis 1a (see column 3 of Table 1 for demographics).

Analysis 1b examined longitudinal changes in EF from 30 to 78 months, but only in the subset of children who had completed their 30-month MEFS assessment before the onset of COVID-19 lockdown restrictions in the United Kingdom. Ninety-four children were included in this analysis. Table 1 shows the demographic characteristics of these children in column 4 (see Table S1 for demographics split by cohort in columns 1 and 2).

In analysis 2a, 2b, and 2c, we examined longitudinal changes in EF for the 57 children in cohort 2 who completed three or more MEFS assessments (at 30, 42, 54, or 78 months). Demographic variables for these children are shown in the fourth column of Table 1. These data are replicated in the fourth column of Table S1 for ease of comparison. Detailed data counts for cohort 2 split by each assessment age are shown in Table S2. The higher amount of missing data at 42 and 54 months was due to the COVID-19 pandemic affecting our ability to collect data at these points in time (Figure 1).

Table 1 Demographic characteristics.

Variable	Cohort one $N=64$	Cohort two $N=75$	Analysis 1a $N=120$	Analysis 1b $N=94$	Analysis 2 $N=57$
Gender					
Boys	33 (52%)	35 (47%)	58 (48%)	49 (52%)	27 (47%)
Girls	31 (48%)	40 (53%)	62 (52%)	45 (48%)	30 (53%)
Maternal education level					
Left school before 16	1 (1.6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
GCSE/O levels or equivalent	4 (6.3%)	10 (13%)	11 (9.2%)	10 (11%)	7 (12%)
A levels or equivalent	6 (9.4%)	10 (13%)	11 (9.2%)	8 (8.5%)	6 (11%)
Trade apprenticeship	0 (0%)	4 (5.3%)	4 (3.3%)	4 (4.3%)	4 (7%)
Some university	5 (7.8%)	6 (8.0%)	9 (7.5%)	7 (7.4%)	4 (7%)
Bachelor's degree	30 (47%)	30 (40%)	55 (46%)	45 (48%)	24 (42%)
Master's degree	12 (19%)	11 (15%)	21 (18%)	14 (15%)	9 (16%)
Doctorate or professional degree	6 (9.4%)	4 (5.3%)	9 (7.5%)	6 (6.4%)	3 (5.3%)
Ethnicity					
White British	54 (84%)	67 (89%)	105 (88%)	84 (89%)	51 (90%)
Asian	1 (1.6%)	0 (0%)	1 (0.8%)	1 (1.1%)	0 (0%)
Black African	0 (0%)	1 (1.3%)	0 (0%)	0 (0%)	0 (0%)
South African	2 (3.1%)	0 (0%)	2 (1.7%)	2 (2.1%)	0 (0%)
White British and South American	2 (3.1%)	0 (0%)	2 (1.7%	0 (0%)	0 (0%)
White British and Asian	1 (1.6%)	2 (2.6%)	3 (2.5%)	2 (2.1%)	2 (3.5%)
White European and Asian	1 (1.6%)	0 (0%)	1 (0.8%)	0 (0%)	0 (0%)
White British and Black African	0 (0%)	2 (2.7%)	1 (0.8%)	1 (1.1%)	1 (1.7%)
White British and Black Caribbean	0 (0%)	2 (2.7%)	2 (1.7%)	2 (2.1%)	2 (3.5%)
White British and Other European	3 (4.7%)	1 (1.3%)	3 (2.5%)	2 (2.1%)	1 (1.7%)

In addition to these five analyses, we also examined the experiences of our participants during the COVID-19 period using a COVID-19 questionnaire. Of the 139 families who participated in this study, 88 completed this COVID-19 questionnaire. To test the relationship between specific COVID-19 related variables and the COVID-19 questionnaire, the dataset of 94 children from analysis 1b was used. Of these 94 families, only 55 completed the COVID-19 questionnaire. Forty-eight out of these 55 families provided a response to every question in the COVID-19 questionnaire.

Procedure

The MEFS (Carlson & Zelazo, 2014) is a 2- to 6-min tablet task (we used a first generation 12.9 inch iPad Pro) based on the DCCS tasks (Zelazo, 2006). The task involves increasingly difficult levels requiring a child to sort cards into a virtual box according to dimensions such as size, shape, or color (see Carlson, 2021 for examples of the MEFS task). Before each level, the child receives a demonstration and rule checks to ensure understanding of the rules for that level. The child was then given a rule to follow based on one of the aforementioned dimensions. The rule was restated on the first two trials, e.g., "If it's a small elephant, it goes in the small elephant box" and the relevant dimension was emphasized, e.g., "Here is a small elephant."

In the next five trials, a prompt was given to ensure the child was ready, e.g., "Get ready!" After five trials, the experimenter announced a "new game" in which the rule was switched. The child was asked to follow this new rule, e.g., "the small elephant goes in the big elephant box." To complete the task, children must focus on the task and pay attention, remember and update the rule/s, inhibit the prepotent response from the previous rule, and engage flexibility to switch rules. An age-appropriate starting level was selected automatically by the app based on the child's age and test norms. Testing continued with a criterion score of 80% at each level. If the criterion score was not met at the starting level, the app automatically regressed levels until a lower level was passed, setting the basal level.

There are a total of seven levels within the MEFS task. In level 1, the child is asked to sort the cards by shape. Level 2 focuses on size. Level 3 involves sorting by both color and shape with separated dimensions. Level 4 involves sorting by shape and color, but with integral dimensions. Level 5 involves the same stimuli, however, now the child is asked to switch between the rules every few trials. Level 6 also involves the same stimuli; however, there is the introduction of a border rule (sort by shape with a black border, sort by color with no border), and there is no longer a verbal prompt from the experimenter. In level 7, this border rule is reversed.

Minnesota Executive Function Scale was completed in the lab from 30 to 54 months of age. The experimenter and child were each sat in a chair at a table. The tablet was placed on the table in front of the child. Parents were in a separate room. If the child requested the parent to be in the room, they sat behind the child so as to not influence the child's responses. At 78 months of age, the MEFS task followed a similar procedure but was conducted inside the child's home. The child was asked to sit on a chair at their dining table and the tablet was located in front of them. If no such table was present in the home, they were asked to sit on a sofa next to the experimenter and the tablet was held on the experimenter's knee facing the child. As all MEFS assessments at 78 months of age were conducted in the home setting, these slight differences between the lab to home setting applied to all children equally. The remaining procedures were identical.

It is important to note that 16 MEFS assessments at 42 months of age and 26 MEFS assessments at 54 months of age were conducted in 2020. Due to the COVID-19 pandemic, many of these assessments were conducted remotely. All guidelines from Reflection Sciences (n.d.) on how to conduct remote assessment were followed. During remote assessment, participants sat facing a laptop with their parent sat behind them. A webcam allowed the experimenter to see the child and parent in full. The experimenter screen shared the iPad screen displaying the MEFS app. The experimenter first checked that both the parent and child could hear them sufficiently. To ensure there were no issues with webcam mirroring, the parent was asked to hold up their left hand, and then their right hand. This allowed the experimenter to confirm the side the child was pointing to if necessary. Parents were instructed "not to help [the child] in anyway, we want to know their skills not yours. Your child may look to you for the answers, and instead of helping them ask them to listen to me." The MEFS task was explained in the same manner as for a non-remote assessment except where the child would be instructed to drag the card into the box, the child was now asked to verbally respond naming the box they wanted the card to be sorted into. If the child was unable to name the box, they were asked to point to the box location, and the parent was asked to name the box (only after a point had taken place).

COVID-19 questionnaire

In the winter of 2022, we added a COVID-19 Questionnaire to the study design to assess the experiences of our cohorts during the pandemic. The questionnaire was sent to families via email, and they were invited to complete the questionnaire via a secure link. The questionnaire asked the following questions: (1) did you have any job related changes during the COVID-19 period? (yes/no); (2) on average, how often was your child in formal childcare (including in-person school) during lockdowns? (0–5 days a week); (3) has your child ever had COVID-19? (never, once, twice, 3 or more times); (4) how many caregivers were present in the child's main home during lockdowns? (number); (5) has the child's caregiver(s) ever had COVID-19 (never, once, twice, 3 or more times).

In addition, we asked caregivers to estimate the activities the child was engaged in during lockdowns: (1) on average, how many days per week did your child do the following activities, and (2) on a typical date the activity was done, how much time did your child spend doing the following activities? The activities listed were: watching TV, playing games on a digital device, playing without an adult, free play with an adult, structured play with an adult (e.g., educational games), share book reading (e.g., being read to or reading to an adult), one to one speaking with an adult, time spent outdoors.

Finally, we included a set of additional questions listed below; however, many families responded "Not Applicable" to these queries so the data were too sparse to assess in detail: (1) compared with prelockdown, what was the child's sleep like? (frequency, length, timing of bedtimes, and naps); (2) how long did your child's period of isolation last for each time they had COVID-19? (<5 days up to >10 days); (3) has your child suffered any long-term effects from having COVID-19? (select from a list of symptoms); (4) has your child been formally diagnosed with long COVID? (yes/no); (5) on average, how many days did caregivers spend working from home during lockdowns? (0–5 days); (6) have the child's caregivers suffered any long-term effects from having COVID (select from a list of symptoms); (7) have the child's caregivers been formally diagnosed with long COVID? (yes/no); (8) have any of your child's siblings had COVID-19? [0 to 3+ times]; (9) have the child's siblings suffered any long-term effects from having COVID-19? (select from a list of symptoms); (10) have the child's siblings been formally diagnosed with long COVID? (yes/no).

Methods of analysis

Scoring was automatically calculated by the MEFS app (MEFS App; Carlson & Zelazo, 2014). A total score (ranging from 0 to 100) was calculated using a proprietary algorithm accounting for accuracy and reaction time. Additional scoring measures from the app include highest level passed (0–7) and a standardized score. We did not analyze the standardized scores as these were based on norms outside of the United Kingdom. Where maternal education level was included in models, it was entered as a scaled and centered numerical variable. Here, a maternal education of "left school at or before 16" was entered as 1, "GCSE/O levels or equivalent" as 2, "A Levels or equivalent" as 3, "Trade Apprenticeship" as 4, "Some University" as 5, "Bachelor's Degree" as 6, "Master's Degree" as 7, and "Doctorate or Professional Degree" as 8. This educational scale was adapted from Saleem and Jan (2020) and has been used in prior work with this sample of children (Fibla et al., 2023).

Preliminary analyses showed that the distribution of total score had long tails at 30, 42, and 78 months. Thus, a Student's *t* distribution was used in all statistical models. This approximates the data distribution more robustly, leading to more normally distributed residuals. Longitudinal models were run using the glmmTMB R package (Brooks et al., 2017). This allowed us to capture the within-subject nature of the data while also using the *t* family distribution. The summary function from the R package (R Core Team, 2021) was used to provide regression coefficients. For significant predictors, the estimated magnitude and direction of the effect are reported. For models with a random intercept, a type III Wald χ^2 test from the car package in R (Fox & Weisberg, 2019) was used to assess the contribution of each parameter in reducing residual deviance of the model.

Total score was scaled (0–1.0) within all models. Across analyses, participants were removed for missing data. Normality was assessed by examining Q–Q plots from the DHARMa R package (Hartig, 2024). The impact of outliers was checked at each stage by comparing models from different families that account for the outliers. In analyses 1a and 1b, we report results from a Student's *t* family that is robust to the effect of outliers, and we report the linear model results in the Supplementary material for comparison. Models for analyses 2a and 2b were run using a mixed-effects model from the beta family to capture the shape that requires asymptotes at 0 and 1. We report results from linear mixed-effects models (Gaussian family) in the Supplementary material for comparison. In no cases were our results affected by outliers.

Analysis 1a examined whether individual differences in EF at 30 months predicted individual differences in EF at 78 months, that is, were individual differences stable longitudinally across this period of development. This was examined in a regression model with total score at 78 months as the dependent variable and total score at 30 months as the predictor variable. A significant effect of total score at 30 months on the total score at 78 months would indicate robust longitudinal stability in EF. This analysis also included two other predictor variables. Maternal education level was included to examine whether this socioeconomic factor modulated EF outcomes at 78 months or interacted with the total score at 30 months, that is, modulated how predictive individual differences at 30 months were of the total score at 78 months. We also included the age at testing at 78 months as this variation undoubtedly impacted the total score (e.g., children tested at 70 months would be expected to have lower total scores than children tested at 86 months).

Analysis 1b followed on from analysis 1a by asking if the longitudinal stability of individual differences from 30 to 78 months was modulated by COVID-19. Here, we restricted the sample to the subset of children who completed their 30-month assessment prior to the first lockdown in the United Kingdom. This gave us an assessment of EF abilities at 30 months of age without any confounding influence of the pandemic. We then examined if individual differences from 30 to 78 months were impacted by COVID-19 by adding "months since COVID" to the regression model. This tested whether EF at 78 months varied systematically depending on when the child was tested relative to the first COVID-19 lockdown (which would be indicated by a main effect of "months since COVID"). This analysis also probed if the longitudinal stability of individual differences from 30 to 78 months varied depending on this COVID-related variable (which would be indicated by an interaction between total score at 30 months and "months since COVID").

Analysis 2a focused on the fine-grained trajectory of MEFS scores from the subset of children from cohort 2 who had at least 3 MEFS assessments from 30 to 42 to 54 to 78 months. Here, we conducted a linear mixed-effects model with a random effect for participant and a random slope over age in months at time of testing. These random effects estimate a linear growth curve for each child with a subject-specific intercept at the first time point and a subject-specific slope over observations. By including fixed effects in this model, we estimated whether group-level trends in total score growth covaried with other factors. We included age in months at time of testing to capture group-level growth in total score (the dependent variable). We also included maternal education to, once again, probe if the group-level trajectory varied based on this socioeconomic factor.

Analysis 2b followed on from analysis 2a, asking if the fine-grained trajectories of MEFS total score were modulated by COVID-19. Here, we added "age at COVID onset" as a predictor of total score. We expected that there might be an interaction between age in months at time of testing (i.e., group-level growth over age) and age at COVID onset. Such an interaction would indicate that the group-level growth curve differed for children who were, for instance, younger at the onset of the pandemic vs. older.

Finally, analysis 2c followed on from analysis 2b by focusing on the individual growth curve estimates for each child, that is, on the random effects from the model. This analysis tested if any effects of age at COVID onset were reflected in an alteration of children's individual trajectories. For this analysis, we began with a growth curve model of MEFS scores over age:

$$\text{TotalScore} \sim \text{Age MEFS} + (1 + \text{Age MEFS} \mid \text{ID})$$

This model was run with a beta family and logit link to capture the shape that has asymptotes at 0 and 1. As this full model with all random effects and allowing for correlation between the random intercept and slope did not converge with the glmmTMB package in R, this model was run in brms, with 4 chains of 10,000 iterations, of which 5,000 were warmup iterations, and thinning every second observation. The model converged with no divergent transitions and Rhat < 1.1.

We then extracted for each participant the coefficient of the intercept and the slope over age and correlated those coefficients with age at COVID onset. If neither parameter correlated, it would demonstrate the effects in 2b might simply be age effects. If the effects correlated with the intercept, it would demonstrate baseline differences before the onset of COVID-19 which would be surprising as all baseline observations were from before the start of the first COVID-19 lockdown. Finally, if the random slope over age correlated with age at COVID onset that would indicate a difference in individual trajectories based on the child's age at COVID onset.

All data and analyses from this study can be found at https://doi.org/10.17605/OSF.IO/X48EN.

Results

Before reporting the results of our primary analyses, it is first useful to describe how our cohort of children experienced COVID-19 and the national lockdowns which took place from 2020 to 2021. Results show 40% of the children within our sample did not receive any kind of formal childcare during the lockdowns; however, 52% did experience 1 day or more in formal childcare (including preschool or primary school). Results also show that 60% of our sample did have the COVID-19 virus once or more (for full details, see Table S3). Families were also asked about the number of days per week their child did different activities. The most popular activities completed >6 days per week on average were time spent outdoors ($M = 6.05$ days), 1–1 speaking with an adult ($M = 6.42$ days), and book reading ($M = 6.0$ days; for full details, see Table S4). Time spent watching TV was a bit lower ($M = 5.37$ days), followed by free play with an adult ($M = 4.74$ days), structured play ($M = 4.05$ days), and finally video games ($M = 3.71$ days).

To test whether maternal education level effects could have been exacerbated during lockdowns due to other socioeconomic variables, a correlation between job changes and maternal education level was examined. There was no correlation found between job changes experienced during lockdown and maternal education level, ($t(52) = 0.61$, $p = .54$). We also

Table 2 Generalized linear mixed model results predicting Minnesota Executive Function Scale (MEFS) total score at 78 months from MEFS at 30 months, maternal education level, and age at test for the 78 months assessment.

Predictor	β	*SE*	*z*	*p*
(Intercept)	0.74***	0.01	84.33	<.001
Total score (30)	0.92***	0.19	4.75	<.001
Maternal education level	0.00	0.01	0.04	.966
Age at test	0.00	0.00	1.77	.077
Total score (30):maternal education level	−0.63**	0.23	−2.72	.006
Total score (30):age at test	−0.03	0.03	−0.99	.322
Maternal education level:age at test	0.00	0.00	0.85	.396
Total score (30):maternal education level:age at test	0.02	0.01	1.73	.083
logLik	120.7			
σ^2	0.0064			

Note. Age at test refers to the age in months at the follow-up (78 months) MEFS assessment. AIC = −223.3, BIC = −198.2. logLik = log-likelihood. σ^2 = Dispersion parameter for Student's *t*. *p*-values: *<.05, **<.01, ***<.001.

examined whether the two key COVID-19 variables examined below (months since COVID, age at COVID onset) correlated with days each child was in formal childcare or whether the child or caregiver had COVID-19. There was a moderate positive relationship between the child's age at COVID onset and the number of times the child had COVID-19 (run using a Poisson distribution: $z(48) = 1.8$, $p = .07$). Thus, children who were older at COVID onset in our sample tended to have COVID-19 more often.

Analysis 1a: do individual differences in EF at 30 months predict individual differences at 78 months?

Analysis 1a examined whether individual differences in MEFS total score at 30 months predict individual differences in MEFS total score at 78 months. We also examined if the longitudinal stability of these individual differences was modulated by maternal education and age at testing at the 78-month assessment point. One hundred and twenty total participants were included in the model, run in R as:

$$\text{glmmTMB(Total score}\left[78\right] \sim \text{Total score}\left[30\right] \\ \times \text{Maternal education level} \times \text{Age at testing)}$$

Minnesota Executive Function Scale total score at 30 months of age significantly predicted MEFS total score at 78 months of age, $\beta = 0.92$, $z = 4.75$, $p < .001$ (see Table 2 and Figure S1 for details). There was also a significant interaction between total score at 30 months and maternal education, $\beta = -0.63$, $z = -2.72$, $p = .006$ (see Table 2; Table S5 shows that this effect was robust to outliers). As is shown in Figure 2, children from families with a more highly educated mother had higher total scores at 30 months and generally higher total scores at 78 months resulting in an overall shallower linear relationship between these measures. By contrast, some children with a less educated mother had lower total scores at both time points, resulting in a steeper linear relationship over development.

Analysis 1b: was the longitudinal stability of individual differences from 30 to 78 months modulated by COVID-19?

Analysis 1b examined whether the longitudinal stability of individual differences in MEFS total score from 30 to 78 months was modulated by variations in the time interval between the onset of the first COVID-19 lockdown and the age at testing. To examine this question, we added "months since COVID" as a predictor to the regression model from analysis 1a. Ninety-four total participants were included in the model, run in R as:

$$\text{glmmTMB(Total score [78]} \sim \text{Total score [30]} \\ \times \text{Maternal education level} \\ \times \text{(Age at testing} + \text{Months since COVID)}$$

Note that "months since COVID" was allowed to interact with total score at 30 months and maternal education level, but not with age at time of testing as the latter variable covaried with months since COVID. In addition, note that when we re-ran analysis 1a on the subset of children included in analysis 1b, we replicated the significant total score at 30 months main effect, $\beta = 1.20$, $z = 4.90$, $p < .001$, as well as the interaction between the total score at 30 months and maternal education, $\beta = -0.48$, $z = -1.98$, $p = .048$.

Table 3 shows the results of analysis 1b. Once again, there was strong evidence that individual differences in EF at 30 months predicted individual differences in EF at 78 months (see the main effect of total score at 30 months). There was evidence of an interaction between total score at 30 months and age at testing, $\beta = 0.88$, $z = 2.51$, $p = .012$, as well as evidence of an interaction between maternal education and age at testing, $\beta = -0.02$, $z = -2.53$ $p = .011$. As shown in Table 3, these two-way interactions were subsumed by a three-way interaction between total score at 30 months, maternal education, and age at testing, $\beta = 0.65$, $z = 3.37$, $p < .001$. Children from families with a more highly educated mother showed a consistent relationship between total score at 30 months and total score at 78 months, regardless of when they were tested (Figure S2). By contrast, there were larger differences in total score at 78 months for the children

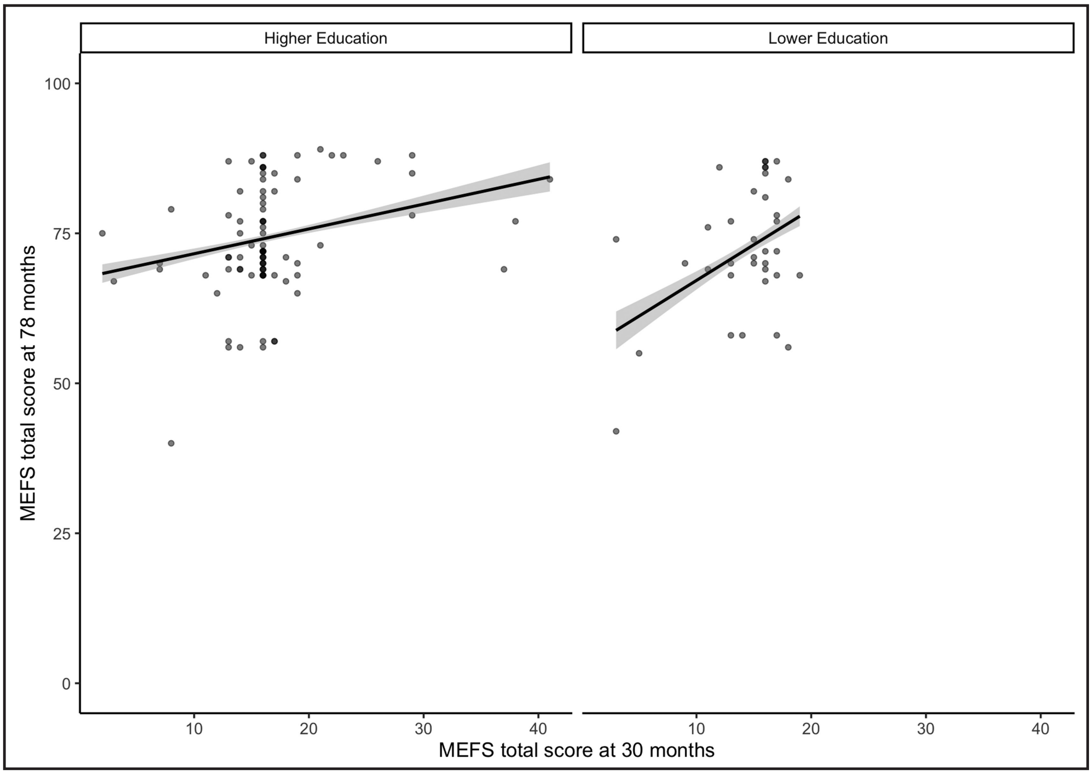

Figure 2 Graph showing Minnesota Executive Function Scale (MEFS) total score at the 78-month assessment by MEFS total score at the 30-month assessment. Left panel shows children from families with a mother who had a bachelor's degree or higher. The right panel shows children from families with a mother with an educational level below a bachelor's degree. Line shows the regression of the model predictions. Error shading indicates the confidence intervals (standard error of the predicted mean). Dots show data values with darker shading indicating overlapping observations across children.

Table 3 Generalized linear mixed model results predicting Minnesota Executive Function Scale (MEFS) total score at 78 months from MEFS at 30 months, maternal education level, age at test (78 months), and months since COVID-19.

Predictor	β	*SE*	*z*	*p*
(Intercept)	0.74***	0.01	65.11	<.001
Total score (30 months)	1.49***	0.27	5.56	<.001
Maternal education level	0.01	0.01	0.75	.453
Age at test	−0.01	0.01	−0.70	.487
Months since COVID-19	−0.00	0.00	−1.16	.248
Total score (30):maternal education level	−0.23	0.26	−0.91	.363
Total score (30):age at test	0.88*	0.35	2.51	.012
Total score (30):months since COVID-19	0.07**	0.03	2.71	.007
Maternal education level:age at test	−0.02*	0.01	−2.53	.011
Maternal education level:months since COVID-19	0.00	0.00	−0.16	.875
Total score (30 months):maternal education level:age at test	0.65***	0.19	3.37	<.001
Total score (30 months):maternal education level:months since COVID-19	−0.04	0.04	−0.94	.346
logLik	97.4			
σ^2	0.0062			

Note. Age at test refers to the age in months at the follow-up (78 months) MEFS assessment. AIC = −168.7, BIC = −135.7, logLik = 97.4, −2 × log(L) = −194.7. σ^2 = Dispersion parameter for Student's *t*. *p*-values: *<.05, **<.01, ***<.001.

from families with a less educated mother based on when they were tested with children younger at test showing the lowest total scores at the 78-month assessment.

Critically, Table 3 shows that there was also a COVID-related finding in analysis 1b: there was a robust interaction between total score at 30 months and the months since COVID-19 predictor, $\beta = 0.07$, $z = 2.71$, $p = .007$. This interaction is shown in Figure 3. Children who were tested in the first 1–3 years after the first COVID-19 lockdown showed a shallower linear relationship between total score at 30 months and total score at 78 months.

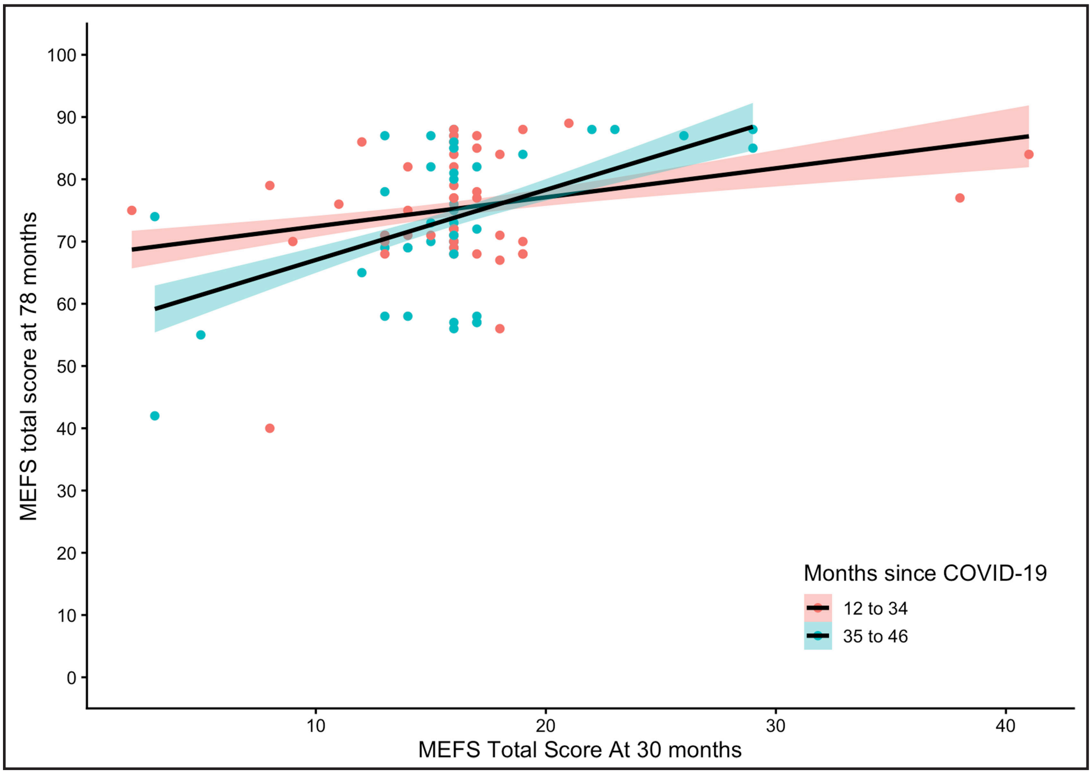

Figure 3 Graph showing Minnesota Executive Function Scale (MEFS) total score at the 78-month assessment by MEFS total score at the 30-month assessment. Red shading shows children who were tested 12–34 months after the first COVID-19 lockdown in the United Kingdom. Blue shading shows children who were tested 35–46 months after the first COVID-19 lockdown. Lines, dot shading, and error shading as in Figure 2.

Thus, individual differences at 30 months were less strongly related to individual differences at 78 months in this group. By contrast, children who were tested 3–4 years after the first COVID-19 lockdown showed a steeper relationship between total score at 30 and 78 months. Note that Table S6 shows that these effects were robust to the influence of outliers.

Analysis 2a: examining the fine-grained group-level trajectory of EF from 30 to 78 months

Analysis 1a demonstrated robust individual differences in EF over time. To explore this further, EF performance was examined in a more fine-grained model over age using only data from cohort 2. Minnesota Executive Function Scale total score was examined in a linear mixed-effects model with a random intercept for participant and a random slope over age at test. In addition, the model included fixed effects of age at test to examine group-level growth in MEFS total score over age as well as maternal education to examine if this group-level growth interacted with maternal education. Fifty-seven participants were included in the model, run in R as:

$$\text{glmmTMB[Total score} \sim \text{Age at test} \times \text{Maternal education level} + (1 + \text{Age at test} || \text{ID})]$$

A Wald χ^2 test demonstrated that both age at test ($X^2(1) = 473.10$, $p < .001$) and maternal education level ($X^2(1) = 5.22$ $p = .022$) significantly explained variance in MEFS total score (see Table 4 for regression coefficients; Table S7 shows that these effects are robust to the presence of outliers). Figure 4 shows the individual data points with individual trajectories of total score growth captured by the gray lines. The solid lines show the linear growth estimates from the random effects structure of the model for children from families with a more highly educated mother, while the dashed lines show the linear growth estimates for children from families with a less educated mother. Note that it would be optimal to estimate nonlinear growth curves given the shape of the data; however, this was not possible given the relatively modest sample size. That said, the linear growth estimates do a decent job of characterizing the individual trajectories over time. The fixed effect of age at test is evident in these data with substantial growth in total score over ages, as is the fixed effect of maternal education with children from families with a more highly educated mother showing consistently higher total scores.

Analysis 2b: did COVID-19 modulate the fine-grained group-level trajectory of EF from 30 to 78 months?

Here, we asked whether MEFS total scores were modulated over age by COVID-19. To examine this question, we added "age at COVID onset" to analysis 2a as a predictor of total score. Fifty-seven participants were included in the model, run in R as:

$$\text{glmmTMB[Total score} \sim \text{Age at test} \times \text{Maternal education level} \times \text{Age at COVID onset} + (1 + \text{Age at test} || \text{ID})]$$

As in analysis 2a, age at test significantly explained variance in MEFS total score, $X^2(1) = 386.82$, $p < .001$. Maternal education also continued to significantly explain variance in MEFS total score, $X^2(1) = 3.93$, $p = .047$ (see Table 5 for regression coefficients;

Table 4 Generalized linear mixed model results for predicting Minnesota Executive Function Scale total score from age at test and maternal education level.

	β	*SE*	*z*	*p*
(Intercept)	−1.18***	0.06	−18.65	<.001
Age at test	0.05***	0.00	21.75	<.001
Maternal education level	0.14*	0.06	2.29	.022
Age at test:maternal education level	0.00	0.00	0.56	.577
Dispersion parameter (β family)	15.2			
Random effects	SD	Variance	Corr	
Participant ID (intercept)	0.189	0.036	–	
Age at test	0.002	0.00	0.00	

Note. p-values: *<.05, **<.01, ***<.001.

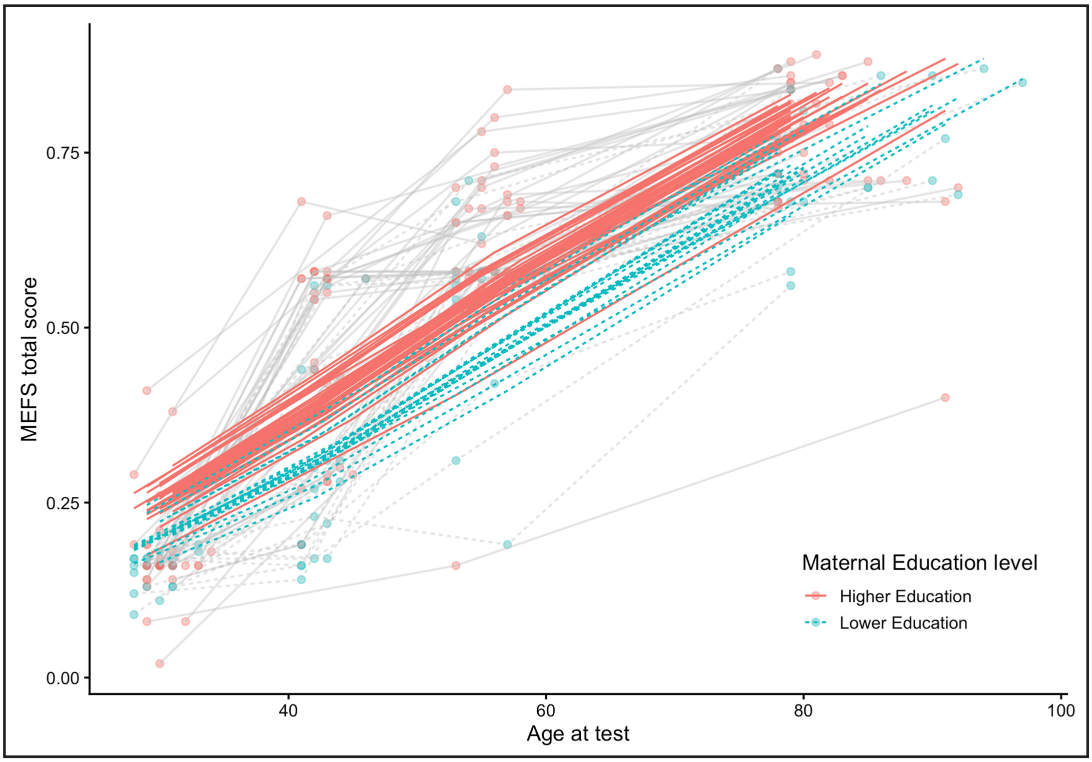

Figure 4 Individual trajectories of Minnesota Executive Function Scale total scores over ages at test for children from cohort 2 with three or more observations. Solid lines connect data points (red dots) for each child longitudinally from families with a more highly educated mother. Solid red lines show predicted linear trajectories for these same children from the linear mixed-effects model. Dashed lines connect data points (blue dots) for each child from families with a less educated mother; dashed blue lines show predicted trajectories for these children.

Table S8 shows that these results were robust to outliers). Most critically, the group-level growth of MEFS total score over age at test interacted with age at COVID onset, $X^2(1) = 5.52$, $p = .019$. This interaction is shown in Figure 5. Note that, rather than show a strict median split in the age at COVID onset measure (which had a median value at 52.2 months), we have opted to split this measure by schooling cohorts with a younger age at COVID onset reflecting children who were in preschool when the first COVID-19 lockdown happened (i.e., <54 months) and an older age at COVID onset reflecting children who were in primary school (i.e., reception) when the first lockdown happened (i.e., ≥54 months). Fifty-four months was used as in the United Kingdom children begin school at the first September after their fourth birthday. As the first lockdown happened 6 months later in March 2020, the youngest children who could be in school at this point would be 54 months of age. As can be seen in Figure 5, children in the preschool cohort showed steeper growth in MEFS total scores than children in the primary school cohort.

Analysis 2c: did COVID-19 modulate fine-grained individual trajectories of EF from 30 to 78 months?

A maximal model to capture the change in MEFS total score over age was fit with brms. To allow for asymptotes at 0 and 100, we rescaled the data to be between 0 and 1 by dividing by 100 and

Table 5 Generalized linear mixed model results for predicting MEFS total score from age at test, age at COVID-19 onset, and maternal education level.

Predictor	β	*SE*	*z*	*p*
(Intercept)	−1.20***	0.07	−16.07	<.001
Age at test	0.05***	0.00	19.67	<.001
Maternal education level	0.15*	0.08	1.98	.047
Age at COVID onset	0.00	0.01	0.21	.831
Age at test × maternal education level	0.00	0.00	0.57	.572
Age at test × age at COVID onset	−0.00*	0.00	−2.35	.019
Maternal education level × Age at COVID onset	−0.00	0.01	−0.21	.836
Age at test × Maternal education level × Age at COVID onset	−0.00	0.00	−0.72	.471
Dispersion parameter (β family)	15.6			
Random effects	SD	Variance	Corr	
Participant ID (intercept)	0.190	0.036	–	
Age at test	0.000	0.000	0.00	

Note. p-values: *<.05, **<.01, ***<.001.

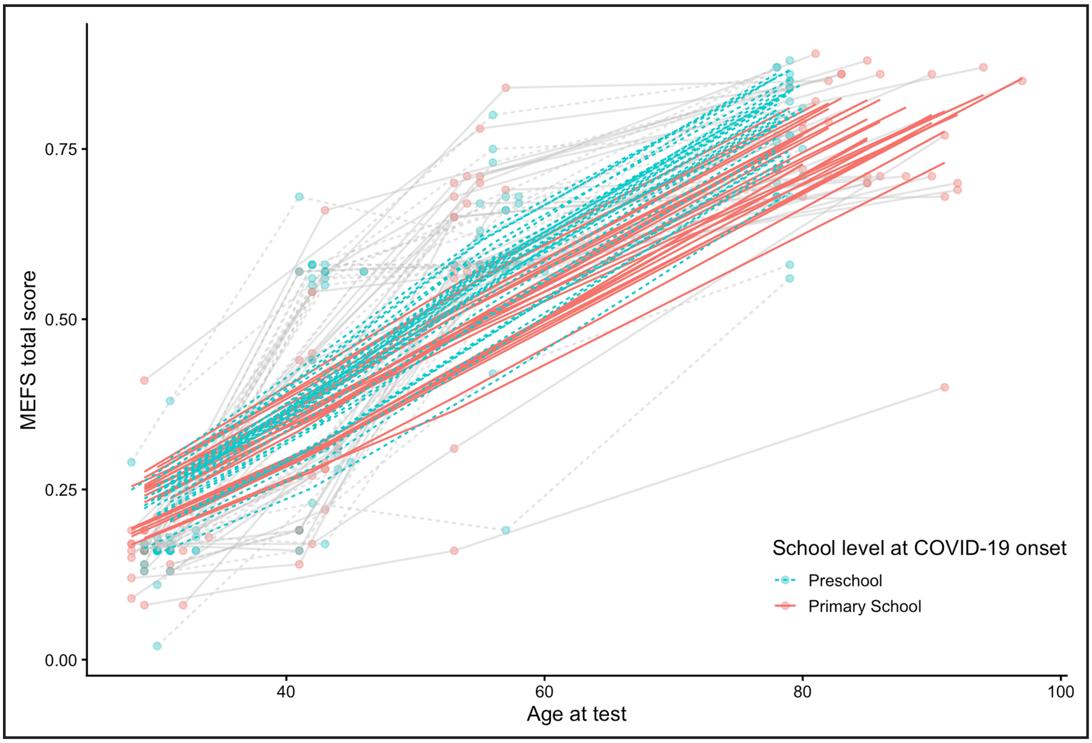

Figure 5 Individual trajectories of Minnesota Executive Function Scale total scores over ages at test for children from cohort 2 with three or more observations. Solid gray lines connect data points (red dots) for each child longitudinally for children who were in primary school during the first COVID lockdown in the United Kingdom. Solid red lines show predicted linear trajectories for these same children from the linear mixed-effects model. Dashed lines connect data points (blue dots) for each child in preschool during the first COVID lockdown; dashed blue lines show predicted trajectories for these children.

fit the model with a Bayesian generalized linear mixed-effects regression model from the beta family. This model took the form:

$$\text{brm}[\text{Total score} \sim \text{Age MEFS} + (1 + \text{Age MEFS} \mid \text{ID})]$$

This model was fit to allow for the full random effect structure, as opposed to the previous analyses that removed the correlation between slope and intercept in the random effect structure due to nonconvergence. Having fitted this model, we extracted the participant-level coefficients for intercept and slope to correlate these coefficients with age at COVID onset. The intercept term represents the participant score at 30 months, and the slope represents their linear growth in total score over time.

There was little evidence of a correlation between the participant intercept and age at COVID onset ($t(55) = -1.94$, $p = .057$); however, there was evidence of a correlation between the participant slope over age and the age at COVID onset ($t(55) = -2.67$,

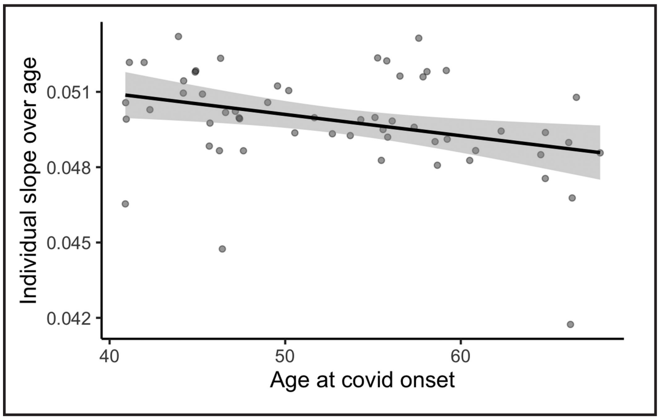

Figure 6 Graph showing the correlation between the participant slope over age from a Bayesian model capturing longitudinal changes in Minnesota Executive Function Scale total scores and age at COVID-19 onset.

$p = .01$). This correlation is shown in Figure 6. These results indicate that the effect of age at COVID onset seen in analysis 2b is not due to baseline differences in ability but rather is an emergent effect that impacted the growth of EF over time such that the preschool children in our sample who were younger at COVID-19 onset showed steeper growth than the children in the first year of primary school.

Discussion

The goal of this study was to explore the longitudinal stability of individual differences in EF from the toddler period through early childhood using the same measure over time. This contrasts with prior longitudinal studies that varied the index of EF across time making it difficult to assess the continuity of EF in early development. As the study was also on-going during a global pandemic, we were well-positioned to assess the impact of the COVID-19 pandemic on early cognitive development. Critically, we had a sample of children who were assessed at the same time point—30 months of age—and this assessment was complete before the onset of the first COVID-19 lockdown.

Results showed an expected increase in EF abilities from 30 to 78 months with robust individual differences in performance over time. Importantly, a fine-grained analysis of cohort two demonstrated a year-on-year increase in EF from 30 to 54 months of age, with a further age-related increase to 78 months of age. These results provide evidence that individual differences in EF measured using the MEFS task are longitudinally stable over a 4-year period from as early as 30 months of age. At each age, age-related developments in EF were reflected by an increase in MEFS total score, and individual differences in the total score were predictive over development across both cohorts of children. These findings support previous research suggesting that EF is longitudinally stable from 4 to 6 years (Helm et al., 2020) and research indicating performance on DCCS-based measures is strongly associated with age (Doebel & Zelazo, 2015).

Previous research suggests socioeconomic factors including maternal education level are important for EF (Hackman et al., 2015; Lawson et al., 2018). Results of the present study support this. In particular, analyses 1a, 1b, 2a, and 2b revealed that maternal education level was consistently related to the MEFS total score, with higher MEFS scores for children with a more highly educated mother. One important factor here may be that mothers with a higher education level themselves have better EF skills and model such skills during interactions with the child. For example, Kao et al. (2018) demonstrated that parents who performed better on EF tasks had children who also performed better on EF tasks. Other studies have found that highly educated parents promote children's cognitive development in a number of ways, including engaging in more stimulating activities, spending more time with their children, and engaging in more complex speech patterns (Landry et al., 2006). Other researchers have suggested that correlations between a parent's education level and child outcomes may be the result of parental characteristics that lead parents to be both good students and good parents (see Duncan & Magnuson, 2012 for a review). Executive functions are also considered to be heritable and related to intelligence (Friedman et al., 2006, 2008). For instance, parental education is associated with child IQ (Lemos et al., 2011), so there may be a genetic component to these maternal education effects. Thus, there are many factors that could explain why a higher maternal education level supports a child's EF. Critically, this interaction has important implications. As discussed by Waters et al. (2021), increased access to educational opportunities for parents, particularly those with a low-level of education, may be an important way to support children's cognitive development.

A central focus of our report was to examine the impact of the COVID-19 pandemic on EF. As a first step in this direction, we used data from a COVID-19 questionnaire that was sent to families to characterize their experiences of the pandemic and the associated national lockdowns. Analyses of these data revealed broad variation in the daily activities of children during lockdowns, as well as variation in which children were in formal childcare during the pandemic. We did not find strong relationships between these questionnaire data and the two COVID-related variables we examined in the primary analyses, although there was some evidence that children who were older at the onset of the first lockdown tended to have COVID-19 more often.

Next, we examined how COVID-19 impacted individual differences in EF trajectories from 30 to 78 months. Analysis 1b revealed that the stability of individual differences in EF from 30 to 78 months varied depending on whether children were tested 1–3 years after the first lockdown vs. 3–4 years after the first lockdown. Interestingly, we found that individual differences were less robust (i.e., a shallower linear slope in the regression model) for children tested 1–3 years after the first lockdown. This was a period marked by major changes in social and schooling routines. Thus, one interpretation is that these disruptions impacted the within-subject stability of individual differences in EF, and these effects were most severe close in time to the pandemic. Consistent with this, previous research indicates a decline in EF was evident as soon as six months after the start of lockdown (Navarro-Soria et al., 2023). Interestingly, however, children with the lowest EF scores at 78 months had low scores at 30 months and were generally tested 3–4 years after the first lockdown. This suggests that low-performing children might have felt the impact of the pandemic for many years after the pandemic was over. A child's ability to overcome the impact of the pandemic on EF may be strongly influenced by a child's preexisting cognitive abilities.

In particular, a strong baseline EF at 30 months of age may have served as a protective factor, supporting children's ability to adapt and recover from the perturbations created by multiple lockdown periods. This interpretation is supported by research indicating children with stronger early EF were more capable of overcoming challenges associated with risky environments (Powell & Gunnar, 2024). It is also likely these children experienced EF recovery—the idea that EF can recover following adversity, particularly when exposed to stable environments (Hostinar et al., 2012). In contrast, children with a lower baseline EF at 30 months of age continued to show lower EF scores at 78 months of age, postpandemic, particularly when tested several years after the first lockdown.

Analyses 2b and 2c showed that age at COVID-19 onset was associated with individual differences in the slope of change over time, with children who were in the first year of primary school during the first lockdown showing a shallower change in EF skills over time. These findings are consistent with results from Perry et al. (2023) who found early EF development was negatively affected by pandemic-related disruptions. While primary school teachers were encouraged to provide at-home support in a number of ways during the COVID-19 pandemic (Moss et al., 2020), this support was mainly focused on math and English, with a reduced focus on personal, social, and emotional development (Nash et al., 2022). Parents also self-reported that their own, and their child's, engagement with these school-provided resources declined as lockdowns went on (Tracey et al., 2022). Consequently, children who were disrupted during primary school and would usually be receiving enriching input in a school setting, had reduced academic, personal, social, and emotional development opportunities. School closures also reduced opportunities for routine socialization with peers in early development. It is known that EF abilities are adversely affected when peer socialization experiences are restricted (Holmes et al., 2016). Note that children in our study who were older at the first COVID-19 lockdown also tended to have the SARS-CoV-2 virus more often. Thus, it is also possible that these periods of illness contributed to the shallower growth in EF skills.

In conclusion, results from the present study reveal robust longitudinal stability in individual differences in EF skills from 30 months of age to 78 months of age. Importantly, this stability was assessed using the same measure over time, reducing the possibility that our findings were impacted by variance in the skills measured at different ages. The MEFS task not only revealed longitudinal stability but also detected shifts in EF skill over time based on differences in maternal education. Most critically, this measure detected the impact of the global pandemic on early cognitive development. Here, we found that COVID-19 disrupted the longitudinal stability of EF during the initial 1–3 years after the first lockdown. In addition, preschool children had steeper growth in EF skills after the first lockdown relative to children in the first year of primary school. It is possible this reflects the importance of peer socialization to early EF skills during primary school such that withdrawal of typical peer interactions during the pandemic slowed the development of EF skills for these children.

Supplementary material

Supplementary material is available at *Child Development* online.

Data availability

The data necessary to reproduce the analyses presented here are publicly accessible. The analytic code necessary to reproduce the analyses presented in this paper is publicly accessible. Data and analytic code can be found at https://doi.org/10.17605/OSF.IO/X48EN. Data are also available from the first author upon reasonable request. The analyses presented here were not preregistered. The materials necessary to attempt to replicate the findings presented here are not publicly accessible.

Author contributions

Eleanor Johns (Conceptualization [supporting], Data curation [equal], Formal analysis [lead], Methodology [supporting], Project administration [supporting], Writing—original draft [lead], Writing—review & editing [lead]), Samuel Forbes (Conceptualization [supporting], Data curation [equal], Formal analysis [equal], Methodology [supporting], Project administration [supporting], Visualization [lead], Writing—original draft [supporting], Writing—review & editing [equal]), Lourdes M. Delgado Reyes (Conceptualization [supporting], Data curation [supporting], Methodology [supporting], Project administration [supporting]), Charlotte Buck (Data curation [supporting], Formal analysis [supporting], Writing—original draft [supporting]), and John P. Spencer (Conceptualization [lead], Data curation [lead], Formal analysis [equal], Funding acquisition [lead], Project administration [lead], Writing—original draft [equal], Writing—review & editing [equal])

Funding

This study was supported by Eunice Kennedy Shriver National Institute of Child Health and Human Development R01HD083287 awarded to J.P.S.

Conflicts of interest

The authors declare no potential conflict of interests.

Acknowledgments

The authors thank the families who participated in this longitudinal research for their robust support over time. Thanks to Laura Marshall, Jo Walther, Nicky Bridgham, Jordan McCarthy, and the rest of the "NIH Team" for help with recruitment, scheduling, and data collection.

References

Ahmed, S. F., Tang, S., Waters, N. E., & Davis-Kean, P. (2019). Executive function and academic achievement: Longitudinal relations from early childhood to adolescence. *Journal of Educational Psychology*, *111*, 446–458. https://doi.org/10.1037/edu0000296

Bell, M. A. (2001). Brain electrical activity associated with cognitive processing during a looking version of the a-not-b task. *Infancy: The Official Journal of the International Society on Infant Studies*, *2*, 311–330. https://doi.org/10.1207/S15327078IN0203_2

Bell, M. A. (2012). A psychobiological perspective on working memory performance at 8 months of age. *Child Development*, *83*, 251–265. https://doi.org/10.1111/j.1467-8624.2011.01684.x

Bernier, A., Carlson, S. M., & Whipple, N. (2010). From external regulation to self-regulation: Early parenting precursors of young children's executive functioning. *Child Development*, *81*, 326–339. https://doi.org/10.1111/j.1467-8624.2009.01397.x

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*, 378–400. https://doi.org/10.32614/RJ-2017-066

Brooks, P. J., Hanauer, J. B., Padowska, B., & Rosman, H. (2003). The role of selective attention in preschoolers' rule use in a novel dimensional card sort. *Cognitive Development*, *18*, 195–215. https://doi.org/10.1016/S0885-2014(03)00020-0

Broomell, A. P., & Bell, M. A. (2022). Longitudinal development of executive function from infancy to late childhood. *Cognitive Development*, *63*, 101229. https://doi.org/10.1016/j.cogdev.2022.101229

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*, 205–228. https://doi.org/10.1080/87565640801982312

Buss, A. T., & Spencer, J. P. (2014). The emergent executive: A dynamic field theory of the development of executive function. *Monographs of the Society for Research in Child Development*, *79*, vii, 1–103. https://doi.org/10.1002/mono.12096

Carlson, S. (2021). Minnesota Executive Function Scale: Technical Report. 1-34. Excelsior, MN: Reflection Sciences, Inc.

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology*, *40*, 1105–1122. https://doi.org/10.1037/0012-1649.40.6.1105

Carlson, S. M., & Zelazo, P. D. (2014). *Minnesota executive function scale: Test manual*. Reflection Sciences: Excelsior.

Cheng, C., Kaldy, Z., & Blaser, E. (2020). Coding of featural information in visual working memory in 2.5-year-old toddlers. *Cognitive Development*, *55*, 100892. https://doi.org/10.1016/j.cogdev.2020.100892

Chichinina, E. A., & Gavrilova, M. N. (2022). Growth of executive functions in preschool-age children during the COVID-19 lockdown: Empirical evidence. *Psychology in Russia: State of the Art*, *15*, 124–136. https://doi.org/10.11621/pir.2022.0209

Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural bases of, inhibitory control in reachinga. *Annals of the New York Academy of Sciences*, *608*, 637–676. https://doi.org/10.1111/j.1749-6632.1990.tb48913.x

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

Diamond, A., Carlson, S. M., & Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology*, *28*, 689–729. https://doi.org/10.1207/s15326942dn2802_7

Doebel, S., & Zelazo, P. D. (2015). A meta-analysis of the dimensional change card sort: Implications for developmental theories and the measurement of executive function in children. *Developmental Review*, *38*, 241–268. https://doi.org/10.1016/j.dr.2015.09.001

Duncan, G. J., & Magnuson, K. (2012). Socioeconomic status and cognitive functioning: Moving from correlation to causation. *Reviews: Cognitive Science*, *3*, 377–386. https://doi.org/10.1002/wcs.1176

Duncan, R. J., McClelland, M. M., & Acock, A. C. (2017). Relations between executive function, behavioral regulation, and achievement: Moderation by family income. *Journal of Applied Developmental Psychology*, *49*, 21–30. https://doi.org/10.1016/j.appdev.2017.01.004

Dunn, J. R. (2010). Health behavior vs the stress of low socioeconomic status and health outcomes. *JAMA*, *303*, 1199–1200. https://doi.org/10.1001/jama.2010.332

Eakin, L., Minde, K., Hechtman, L., Ochs, E., Krane, E., Bouffard, R., & Looper, K. (2004). The marital and family functioning of adults with adhd and their spouses. *Journal of Attention Disorders*, *8*, 1–10. https://doi.org/10.1177/108705470400800101

Fibla, L., Forbes, S. H., McCarthy, J., Mee, K., Magnotta, V., Deoni, S., & Spencer, J. P. (2023). Language exposure and brain myelination in early development. *Journal of Neuroscience*, *43*, 4279–4290. https://doi.org/10.1523/JNEUROSCI.1034-22.2023

Fox, J., & Weisberg, S. (2019). An R companion to applied regression, 3rd ed. Sage. https://www.john-fox.ca/Companion/

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

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*, 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*, 483–527. https://doi.org/10.1016/0885-2014(95)90024-1

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology*, *50*, 1698–1709. https://doi.org/10.1037/a0036633

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

Garon, N., Smith, I. M., & Bryson, S. E. (2014). A novel executive function battery for preschoolers: Sensitivity to age differences. *Child Neuropsychology*, *20*, 713–736. https://doi.org/10.1080/09297049.2013.857650

Gathercole, S. E., Pickering, S. J., & Knight, C. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *18*, 1–16. https://doi.org/10.1002/acp.934

Hackman, D. A., Gallop, R., Evans, G. W., & Farah, M. J. (2015). Socioeconomic status and executive function: Developmental trajectories and mediation. *Developmental Science*, *18*, 686–702. https://doi.org/10.1111/desc.12246

Hartig, F. (2024). Harma: Residual diagnostics for hierarchical (multilevel/mixed) regression models. *R package version 0.4.7*. https://github.com/florianhartig/dharma (Accessed on 4 November 2025).

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*, 411–482. https://doi.org/10.1086/504455

Helm, A. F., McCormick, S. A., Deater-Deckard, K., Smith, C. L., Calkins, S. D., & Bell, M. A. (2020). Parenting and children's executive function stability across the transition to school. *Infant and Child Development*, *29*, e2171. https://doi.org/10.1002/icd.2171

Holmes, C. J., Kim-Spoon, J., & Deater-Deckard, K. (2016). Linking executive function and peer problems from early childhood through middle adolescence. *Journal of Abnormal Child Psychology*, *44*, 31–42. https://doi.org/10.1007/s10802-015-0044-5

Hostinar, C. E., Stellern, S. A., Schaefer, S. M., Carlson, C., & Gunnar, M. R. (2012). Associations between early life adversity and executive function in children adopted internationally from orphanages. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 17208–17212. https://doi.org/10.1073/pnas.1121246109

Howard, S. J., Okely, A. D., & Ellis, Y. G. (2015). Evaluation of a differentiation model of preschoolers' executive functions. *Frontiers in Psychology*, *6*, 285. https://doi.org/10.3389/fpsyg.2015.00285

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, *35*, 20–36. https://doi.org/10.1080/87565640903325691

Johansson, M., Marciszko, C., Brocki, K., & Bohlin, G. (2012). Individual differences in early executive functions: A longitudinal study from 12 to 36 months. *Infant and Child Development*, *25*, 533–549. https://doi.org/10.1002/icd.1952

Kao, K., Nayak, S., Doan, S. N., & Tarullo, A. R. (2018). Relations between parent EF and child EF: The role of socioeconomic status and parenting on executive functioning in early childhood. *Translational Issues in Psychological Science*, *4*, 122–467. https://doi.org/10.1037/tps0000154

Korzeniowski, C. (2023). Impact of COVID-19 pandemic on the development of children's executive functions: Implications for schoolbased interventions. *Journal of Biomedical Research*, *4*, 1120–1135. https://doi.org/10.37871/jbres1776

Landry, S. H., Smith, K. E., & Swank, P. R. (2006). Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Developmental Psychology*, *42*, 627–642. https://doi.org/10.1037/0012-1649.42.4.627

Lawson, G. M., Hook, C. J., & Farah, M. J. (2018). A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Developmental Science*, *21*, 10.1111/desc.12529. https://doi.org/10.1111/desc.12529

Lemos, G. C., Almeida, L. S., & Colom, R. (2011). Intelligence of adolescents is related to their parents' educational level but not to family income. *Personality and Individual Differences*, *50*, 1062–1067. https://doi.org/10.1016/j.paid.2011.01.025

Lowery, K., Nikam, B., & Buss, A. T. (2022). Dimensional label learning contributes to the development of executive functions. *Scientific Reports*, *12*, 11008. https://doi.org/10.1038/s41598-022-14761-2

McCraw, A., Sullivan, J., Lowery, K., Eddings, R., Heim, H. R., & Buss, A. T. (2024). Dynamic field theory of executive function: Identifying early neurocognitive markers. *Monographs of the Society for Research in Child Development*, *89*, 7–109. https://doi.org/10.1111/mono.12478

Miller, S. E., & Marcovitch, S. (2015). Examining executive function in the second year of life: Coherence, stability, and relations to joint attention and language. *Developmental Psychology*, *51*, 101–114. https://doi.org/10.1037/a0038359

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. https://doi.org/10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., & Sears, M. R. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 2693–2698. https://doi.org/10.1073/pnas.1010076108

Moss, G., Allen, R., Bradbury, A., Duncan, S., Harmey, S., & Levy, R. (2020). *Primary teachers' experience of the COVID-19 lockdown—eight key messages for policymakers going forward*. UCL Institute of Education.

Nash, H., Clarke, P., Davies, C., Hart, P., Homer, M., & Mathieson, R. (2022). Progress of children through reception and year 1 during COVID-19 school disruption. https://eprints.whiterose.ac.uk/183418/1/ickle-fr2-compressed.pdf (Accessed on 5 November 2025).

Navarro-Soria, I., Costa-López, B., Collado-Valero, J. A., Juárez-Ruiz de Mier, R., & Lavigne-Cervan, R. (2023). Anxiety, sleep habits and executive function during the COVID-19 pandemic through parents' perception: A longitudinal study. *Psicología, Reflexão e Crítica*, *36*, 8. https://doi.org/10.1186/s41155-023-00251-5

Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the dimensional change card sorting task? *Infant and Child Development*, *11*, 93–105. https://doi.org/10.1002/icd.299

Perone, S., Molitor, S., Buss, A., Spencer, J., & Samuelson, L. (2015). Enhancing the executive functions of 3-year-olds in the dimensional change card sort task. *Child Development*, *86*, 812–827. https://doi.org/10.1111/cdev.12330

Perry, K. J., Perhamus, G. R., Lent, M. C., Murray-Close, D., & Ostrov, J. M. (2023). The COVID-19 pandemic and measurement of preschoolers' executive functions. *Psychological Assessment*, *35*, 986–999. https://doi.org/10.1037/pas0001250

Piaget, J. (1954). *The construction of reality in the child*. Routledge.

Polizzi, C., Burgio, S., Lavanco, G., & Alesi, M. (2021). Parental distress and perception of children's executive functioning after the first COVID-19 lockdown in Italy. *Journal of Clinical Medicine*, *10*, 4170. https://doi.org/10.3390/jcm10184170

Powell, T., & Gunnar, M. R. (2024). *Protective role of executive function skills in high-risk environments*. University of Minnesota.

Prime Minister's Office. (23 March 2020). Prime minister's statement on coronavirus (covid-19): 23 march 2020. https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-23-march-2020 (Accessed on 31 July 2025).

Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., & Rahman, A. (2007). No health without mental health. *The Lancet*, *370*, 859–877. https://doi.org/10.1016/S0140-6736(07)61238-0

R Core Team. (2021). *Tr: A language and environment for statistical computing.* R Foundation for Statistical Computing.

Richmond-Rakerd, L. S., Caspi, A., Ambler, A., d'Arbeloff, T., de Bruine, M., Elliott, M., Harrington, H., Hogan, S., Houts, R. M., Ireland, D., Keenan, R., Knodt, A. R., Melzer, T. R., Park, S., Poulton, R., Ramrakha, S., Rasmussen, L. J. H., Sack, E., Schmidt, A. T., … Moffitt, T. E. (2021). Childhood self-control forecasts the pace of midlife aging and preparedness for old age. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, e2010211118. https://doi.org/10.1073/pnas.2010211118

Saleem, S. M., & Jan, S. S. (2020). Modified Kuppuswamy socioeconomic scale updated for the year 2020. *Indian Journal of Forensic and Community Medicine*, *6*, 1–3. https://doi.org/10.18231/2394-6776.2019.0001

Spencer, J. P., Buss, A. T., McCraw, A. R., Johns, E., & Samuelson, L. K. (2025). Integrating attention, working memory, and word learning in a dynamic field theory of executive function development: Moving beyond the 'component' view of executive function. *Developmental Review: DR*, *75*, 101182. https://doi.org/10.1016/j.dr.2024.101182

Steelandt, S., Thierry, B., Broihanne, M. H., & Dufour, V. (2012). The ability of children to delay gratification in an exchange task. *Cognition*, *122*, 416–425. https://doi.org/10.1016/j.cognition.2011.11.009

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *The Behavioral and Brain Sciences*, *24*, 1–34. https://doi.org/10.1017/S0140525X01003910

Tracey, L., Bowyer-Crane, C., Bonetti, S., Nielsen, D., D'Apice, K., & Compton, S. (2022). *The impact of the COVID-19 pandemic on children's socio-emotional wellbeing and attainment during the reception year. Research report*. Education Endowment Foundation.

Veraksa, A. N., Veresov, N. N., Sukhikh, V. L., Gavrilova, M. N., & Plotnikova, V. A. (2024). Play to foster children's executive function skills: Exploring short- and long-term effects of digital and traditional types of play. *Journal of Early Childhood*, *56*, 687–709. https://doi.org/10.1007/s13158-023-00377-8

Waters, N. E., Ahmed, S. F., Tang, S., Morrison, F. J., & Davis-Kean, P. E. (2021). Pathways from socioeconomic status to early academic achievement: The role of specific executive functions. *Early Childhood Research Quarterly*, *54*, 321–331. https://doi.org/10.1016/j.ecresq.2020.09.008

Wiebe, S. A., Lukowski, A. F., & Bauer, P. J. (2010). Sequence imitation and reaching measures of executive control: A longitudinal examination in the second year of life. *Developmental Neuropsychology*, *35*, 522–538. https://doi.org/10.1080/87565641.2010.494751

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-yearolds. *Journal of Experimental Child Psychology*, *108*, 436–452. https://doi.org/10.1016/j.jecp.2010.08.008

Wijeakumar, S., Ambrose, J., Spencer, J., & Curtu, R. (2017). Modelbased functional neuroimaging using dynamic neural fields: An integrative cognitive neuroscience approach. *Journal of Mathematical Psychology*, *76*, 212–235. https://doi.org/10.1016/j.jmp.2016.11.002

Willoughby, M. T., Wirth, R. J., Blair, C. B.; & Family Life Project Investigators. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, *24*, 418–431. https://doi.org/10.1037/a0025779

Zelazo, P. D. (2006). The dimensional change card sort (dccs): A method of assessing executive function in children. *Nature Protocols*, *1*, 297–301. https://doi.org/10.1038/nprot.2006.46

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. nih toolbox cognition battery (cb): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*, 16–33. https://doi.org/10.1111/mono.12032