Journal of Experimental Child Psychology 235 (2023) 105715

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Executive functioning skills and (low) math achievement in primary and secondary school

Valentín Iglesias-Sarmiento [a], Nuria Carriedo [b,*], Odir A. Rodríguez-Villagra [c,d], Leire Pérez [a]

[a] Department of Evolutionary Psychology and Communication, University of Vigo, Campus Universitario de Ourense, 32004, Vigo, Spain
[b] Departamento de Psicología Evolutiva y de la Educación, National Distance Education University (UNED), 28040 Madrid, Spain
[c] Institute for Psychological Research, University of Costa Rica, San José, 11501-2060, Costa Rica
[d] Neuroscience Research Center, University of Costa Rica, San José 11501-2060, Costa Rica

ARTICLE INFO

Article history:
Received 27 October 2022
Revised 12 May 2023

Keywords:
Learning difficulties
Executive functioning
Mathematical achievement
Processing speed
Fluid intelligence

ABSTRACT

Schoolchildren with better executive functioning skills achieve better mathematics results. It is less clear how inhibition, cognitive flexibility, and working memory combine to predict mathematics achievement and difficulty throughout primary and secondary school. This study aimed to find the best combination of executive function measures for predicting mathematical achievement in Grades 2, 6, and 10 and to test whether this combination predicts the probability of having mathematical difficulties across school grades even when fluid intelligence and processing speed were included in the models. A total of 426 students—141 2nd graders (72 girls), 143 6th graders (72 girls), and 142 10th graders (79 girls)—were cross-sectionally assessed with 12 executive tasks, one standardized mathematical task, and a standardized test of intelligence. Bayesian regression analyses found various combinations of executive predictors of mathematical achievement for each school grade spanning Grade 2 to measures of cognitive inhibition (negative priming) and cognitive flexibility (verbal fluency); Grade 6 to measures of inhibition: resistance to distractor interference (receptive attention), cognitive flexibility (local–global), and working memory (counting span); and Grade 10 to measures of inhibition: resistance to distractor interference (receptive attention) and prepotent response inhibition (stop signal) and working memory (reading span). Logistic regression showed that the executive

* Corresponding author.
E-mail address: ncarriedo@psi.uned.es (N. Carriedo).

https://doi.org/10.1016/j.jecp.2023.105715
0022-0965/

models derived from the Bayesian analyses had a similar ability to classify students with mathematical difficulty and their peers with typical achievement to broader cognitive models that included fluid intelligence and processing speed. Measures of processing speed, cognitive flexibility (local–global), and prepotent response inhibition (stop signal) were the main risk factors in Grades 2, 6, and 10, respectively. Cognitive flexibility (verbal fluency) in Grade 2 and fluid intelligence, which was more stable in all three grades, acted as protective factors against mathematical difficulty. These findings inform practical considerations for establishing preventive and intervention proposals.

Introduction

Executive functioning is considered a set of attentional control top-down processes that allow adaptative behavior in many domains of life, including academia (Zelazo & Carlson, 2020). Although there is some dispute about its conceptualization and the processes considered executive, there is some consensus that inhibition, updating, and shifting are among the most critical (Miyake et al., 2000).

Inhibition is not a unitary construct. Some previous developmental taxonomies and conceptual distinctions (Dempster, 1993; Diamond, 2013; Friedman & Miyake; 2004; Harnishfeger, 1995; Hasher et al., 1999; Nigg, 2000) have distinguished three types of functional and developmentally different inhibitory functions: (a) *resistance to distractor interference,* that is, the ability to resist the distraction of external stimuli not relevant for the current task; (b) *prepotent response inhibition,* that is, the capacity to stop the automatic responses voluntarily; and (c) *cognitive inhibition,* that is, the ability to avoid interference from information that has previously been relevant to the performance of the task but is no longer relevant.

Shifting, also called mental set shifting, mental flexibility, or cognitive flexibility (Diamond, 2013), is defined as the capacity to switch attention flexibly between aspects of a task—responses, rules, criteria, mental representations, stimuli, and so on (Monsell, 2003). Hereafter, we use *cognitive flexibility* to refer to low-level perceptual shifting and conceptual shifting (Glass et al., 2013).

Updating is "the ability to modify the current state of the schema representation in memory to accommodate new entries" (Morris & Jones, 1990, p. 112). It involves different subprocesses: the maintenance and transformation of information in working memory as well as the deletion of irrelevant old content by new content based on varying task demands (Ecker et al., 2014; Kessler & Meiran, 2008). Given that updating measures correlate highly with working memory capacity measures, the two terms could be used interchangeably (Schmiedek et al., 2009), so hereafter we use the term *working memory* (or WM) in line with the majority of previous studies (Packwood et al., 2011).

A considerable amount of the research conducted over the last 20 years has analyzed the relationship among these three main executive functioning skills (hereafter EFs) and academic achievement in mathematics during the school period (see Agostini et al., 2022, and Friso-van den Bos et al., 2013, for reviews). However, as we see below, few of these studies addressed the three EFs conjunctively and their predictive value on mathematical difficulties while also considering other related cognitive constructs such as fluid intelligence and processing speed. In this study, we tried to fill this gap through two analytic strategies: (a) finding a better pattern of the relationships among these EFs when students perform school mathematics tasks across Grades 2, 6, and 10 and (b) classifying students according to their mathematical achievement and identifying the risk and protective factors that are related to students' mathematical difficulties during their development. Both seem to be essential for developing precise educational actions.

EFs and academic achievement in mathematics

Recent studies have found evidence linking EFs to individual differences in mathematical achievement across schooling, although the results are controversial mainly due to the heterogeneity of the mathematical and EF measures used (see Friso-van den Bos et al., 2013, and Nunes de Santana et al., 2022, for reviews).

Conceptually, inhibition has been associated with the need to comprehend mathematical tasks by inhibiting students' first impulse to solve mathematical problems by preventing the application of learned strategies that might not be suitable for the current task context (Bull & Scerif, 2001) and suppressing irrelevant information during task resolution (Passolunghi & Siegel, 2001). Although the evidence is inconclusive, some studies have found significant contributions of response inhibition to mathematical achievement in late primary school grades (e.g., Gerst et al., 2015; Lubin et al., 2016) and, to a lesser extent, during secondary school (Latzman et al., 2010; but see Cragg et al., 2017).

Cognitive flexibility should allow for switching among notations, operations, strategies, and (crucially) steps to be taken in complex mathematical tasks (Archambeau & Gevers, 2018). Two recent meta-analyses have moderately related cognitive flexibility to mathematical achievement (Nunes de Santana et al., 2022; Yeniad et al., 2013), although there is some evidence to the contrary (e.g., Lee et al., 2009; Lubin et al., 2016) that could be associated with the differential complexity of the mathematical tasks (Friso-van den Bos et al., 2013) and/or of the cognitive flexibility tasks (Yeniad et al., 2013) employed.

The involvement of WM in mathematics has been linked to more demanding multiple-step tasks, whose resolution implies the storage and active maintenance of crucial information and the intermediate results obtained (Bull & Lee, 2014). WM has become the cognitive benchmark of mathematical development given that the preschool stage, and its relationship to mathematical achievement appears to hold through primary and secondary school (Cragg et al., 2017; Lee & Bull, 2016). This relationship seems to be moderated by the type of mathematical measure used (higher for global and curricular measures) and the sample used (higher for children with mathematical difficulties) (Friso-van den Bos et al., 2013).

Unfortunately, very few studies have looked at multiple measures of the three EFs concerning math achievement (e.g., Lee & Bull, 2016; van der Ven et al., 2012), and only two studies (to our knowledge) have analyzed how the relationships between EFs and mathematical achievement change with age across primary and secondary school (Cragg et al., 2017; Lee & Bull, 2016) but with a different analytic approach from that adopted in the current study. Both studies reported only stable relationships between WM and curricular achievement in mathematics. Neither study reported significant relationships between the measures of inhibition and cognitive flexibility when they were analyzed with WM. However, methodological and analytical issues must be considered when interpreting these results.

Lee and Bull (2016) focused their influential work on the influence of WM —viewed as a latent variable composed of a verbal memory span task, another visuospatial span task, and a pictorial updating task—on math performance. Only an inhibition/switch factor—composed of three measures of inhibition (one measure of resistance to distractor interference and two measures of prepotent response inhibition) and three perceptual cognitive flexibility measures—were used as a control in a correlational analysis. The authors found significant correlations between WM and math in all grades (1–9), but the correlations between inhibition/switch factor and math did not reach significance at any grade when WM, inhibition, and shifting were considered simultaneously. However, it must be mentioned that the inhibition/switch factor was composed of different experimental conditions of the same tasks.

Cragg et al. (2017) conducted a single hierarchical regression analysis for 8- and 9-year-olds, 11- and 12-year-olds, 13- and 14-year-olds, and young adults. Two contrasts with dummy variables were used to analyze changes in age. The authors used regression analysis to test the relative contributions of inhibition, working memory, and cognitive flexibility to mathematics achievement. They used five executive measures: a verbal span task, a visuospatial working memory task, two prepotent response inhibition tasks with numerical and non-numerical content, and a single cognitive flexibility task. They also used two measures of short-term memory. They found that the total scores for verbal and visuospatial working memory were unique independent predictors of mathematics achievement.

EFs and mathematical difficulties

The influence of executive deficits on students with mathematical difficulties (Agostini et al., 2022; Peng et al., 2018) is not well established in the literature, where only a few studies have analyzed the joint contribution of EFs consistently using various paradigms and with representative samples.

Deficits in prepotent response inhibition have been reported in students with mathematical difficulties in late primary school grades (Szucs et al., 2013) and heterogeneous samples extending into secondary school (Cai et al., 2013; Willcutt et al., 2013). In addition, although the results are controversial, possibly due to different selection criteria and the analytical approach used, differences between students with and without mathematical difficulties have been found during late primary school grades in cognitive inhibition (Passolunghi et al., 1999; Passolunghi & Siegel, 2004; but see Ng et al., 2017) and in resistance to distractor interference in adolescents (Cai et al., 2013; Deng et al., 2022; but see Censabella & Noël, 2007).

Peng et al. (2018) also pointed to deficits in cognitive flexibility as characteristic of students with mathematical difficulties, mainly in primary school groups selected from arithmetic tasks, although to a lesser extent than the other EFs. Specifically, deficits in cognitive flexibility have been reported in Grade 4 students (McLean & Hitch, 1999; Szucs et al., 2013) and in heterogeneous age-equivalent samples in Grades 2 to 5 (McDonald & Berg, 2017) and Grades 4 and 5 (van der Sluis et al., 2004). Andersson (2010) also reported deficits in cognitive flexibility in a longitudinal study with students in Grades 3 and 4 with mixed arithmetic and reading difficulties.

EFs, intelligence, processing speed, and mathematical difficulties

Fluid intelligence and processing speed are two important variables to consider because of their relationship to EFs (Camos & Barrouillet, 2011; Gray et al., 2017) and learning difficulties (e.g., Andersson, 2010; Murphy et al., 2007). Developmentally, both constructs have a similar tendency to that of EFs (e.g., McAuley & White, 2011; Uka et al., 2019), improving more intensely during childhood and more gradually during adolescence (Fry & Hale, 2000).

The role of intelligence in mathematical difficulties is not clearly defined because most studies have excluded students with low IQs in their initial screenings (see Agostini et al., 2022, for a review). Nevertheless, some studies have reported differences between students with mathematical difficulties and those with typical mathematical achievement on verbal and nonverbal intelligence tasks (e.g., Andersson, 2010; Geary et al., 2020; but see Andersson, 2008) and that it was a key factor in classifying students in terms of mathematical achievement (Geary et al., 2020). In addition, during the early primary grades, the importance of processing speed in the acquisition of arithmetic skills has been highlighted (e.g., Bull & Johnston, 1997; Fuchs et al., 2008) and the low processing speed of students with mathematical difficulties has been repeatedly reported (e.g., Cirino et al., 2015; Vukovic & Siegel, 2010; but see Chan & Ho, 2010).

The current study

In this context, this study aimed to analyze the contributions of EFs to academic achievement in mathematics in Grades 2, 6, and 10. The age groups corresponding to these grades were selected due to their consideration as critical periods for EF development (e.g., Epstein, 2001; Huizinga et al., 2006). From a mathematical point of view, to get as close as possible to the educational context, we chose a globally standardized test that covers the main mathematical topics acquired throughout schooling.

The study's first goal was to analyze the extent to which individual differences in EFs were associated with differences in mathematical achievement to identify the best combination of executive predictors for each school grade and to quantify their relative importance. In an attempt to provide new information in this study, we used a Bayesian individual differences approach (Rouder & Morey, 2015) around 12 tasks that the literature has associated with different facets of EF. Based on the differential demands of mathematical tasks associated with each school grade, we expected EFs to contribute to mathematical achievement differentially across the selected grades. Specifically, we predicted that the contributions of inhibition (Gerst et al., 2015; Latzman et al., 2010) and WM (Cragg et al., 2017; Lee & Bull, 2016) would be significant throughout school grades because they continue to develop far beyond adolescence (Carriedo et al., 2016).

We also expected that cognitive flexibility would have a greater influence on primary school students than on secondary school students (Best et al., 2011; Friso-van den Bos et al., 2013) because although adolescents may have more efficient cognitive flexibility than children (Davidson et al., 2006; Garon et al., 2008), they can use overlearned automatic strategies to solve procedural mathematical problems that do not require cognitive flexibility.

Our subsequent goal was to investigate the ability of EFs to classify students based on their prior mathematical achievement (mathematical difficulties vs. typical mathematical achievement) by reference to broader cognitive models, including related variables such as fluid intelligence and processing speed. Based on the results of previous studies (e.g., Agostini et al., 2022; McAuley & White, 2011; Peng et al., 2018; Uka et al., 2019), we expected the models based on EFs to exhibit adequate sensitivity, specificity, and efficiency in this context. Specifically, we expected that response inhibition would emerge as a predictor of mathematical difficulties across the three grades included in this study (e.g., Szucs et al., 2013; Willcutt et al., 2013) given that it continues to develop far beyond adolescence (Carriedo et al., 2016). We also anticipated cognitive flexibility to contribute to the task of discriminating between groups mainly with respect to students in Grades 2 and 6 (McDonald & Berg, 2017; McLean & Hitch, 1999) because it builds over inhibition and working memory and therefore it develops later (Davidson et al., 2006; Garon et al., 2008). Finally, we expected that the contributions of WM would extend across the three grades (Peng et al., 2018) and could be related to verbal tasks to a greater extent (Cai et al., 2013; Cirino et al., 2015).

Furthermore, we predicted that higher fluid intelligence could act as a protective factor against mathematical difficulties across the three grades (Geary et al., 2020). In addition, slower processing speed could constitute a risk factor for presenting mathematical difficulties in situations where the selection tasks involve arithmetic, as in Grade 2 (e.g., Cirino et al., 2015; Vukovic & Siegel, 2010).

Method

Participants

The initial sample for this study included 450 students in Grades 2, 6, and 10 in six Spanish-language primary and secondary schools in Spain. These schools serve middle-class urban and semi-urban areas. According to the information provided by the education professionals, none of the children had sensory deficits or developmental disorders or was at risk for social exclusion. An additional criterion for including students was that their scores on fluid intelligence were above Grade V (intellectual deficit) on the SPM scale of Raven's Standard Progressive Matrices (Raven et al., 1996). In total, 3 students were excluded from the study because they did not complete all the tasks, and 21 students were excluded because they did not meet the intelligence criterion. The final sample comprised a total of 426 children: 141 2nd graders (72 girls), 143 6th graders (72 girls), and 142 10th graders (79 girls).

To identify the groups with difficulties in mathematics, the 25th percentile of the numerical scale of the Battery of Differential and General Mental Aptitudes (BADyG) in its versions Level E1 (Grade 2; Yuste & Yuste, 2011), Level E3 (Grade 6; Yuste et al., 2011), and Level M (Grade 10; Yuste et al., 2012) was used as a cutoff criterion (e.g., Cirino et al., 2015; Geary et al., 2020). Students scoring below the 25th percentile were assigned to the mathematical difficulties group: 18 for Grade 2 (12.77%; M_{age}[1] = 91.61 months, SD = 3.79), 42 for Grade 6 (29.37%; M_{age} = 136.97 months, SD = 5.15), and 28 for Grade 10 (19.72%; M_{age} = 185.68 months, SD = 4.85). The remaining students were assigned to the control group with typical mathematical achievement: 123 for Grade 2 (87.23%; M_{age} = 93.0 months, SD = 3.51), 101 for Grade 6 (70.63%; M_{age} = 137.46 months, SD = 4.29), and 114 for Grade 10 (80.28%; M_{age} = 185.59 months, SD = 4.25).

[1] Age is in months.

Materials

Mathematical achievement

The numerical scale of the BADyG was used to assess mathematical achievement. This scale is organized around two groups of tasks: numerical and problem-solving. The first mainly assesses conceptual and procedural arithmetic knowledge through arithmetic operations of addition and subtraction with two or three operands for Grade 2 and with numerical series of increasing difficulty for Grades 6 and 10. The problem-solving tasks involve various levels of mathematical knowledge. In Grade 2, they are linked to four types of word problems (change, combine, equalize, and compare) with various levels of complexity. In Grades 6 and 10, fractions, decimals, proportions, geometry, algebra, and word problems are included. The final mathematical achievement score comprises the number of correct answers on the numerical scale. In addition, the centile score of the scale was used to select the groups. The reliability indexes provided by the test for the types of tasks (numerical and problem-solving) were .93 and .87 for Grade 2, .87 and .85 for Grade 6, and .93 and .86 for Grade 10, respectively.

Executive measures

We assessed selected EFs based on an in-depth review of previous developmental research that had been empirically validated by confirmatory factor analysis (see Karr et al., 2018, for a review); thus, the selected measures are comparable with those used in previous comparative and predictive studies. The tasks were selected to measure different components of each EF: inhibition (resistance to distractor interference, prepotent response inhibition, and cognitive inhibition), working memory (span and updating), and cognitive flexibility (perceptual and conceptual shifting). The tasks tapped different stimulus modalities (verbal and visuospatial). The tasks did not include mathematical or numerical content (except for the counting span task, in which participants needed to count) that could be related to the criterion variable to avoid overlaps between the predictors and outcome variables and thus ensure that the predictors did not have domain-specific content. The same tasks were used for all three grades except for the receptive attention task (pictures were used for 2nd graders, and letters were used for 6th and 10th graders) to ensure that differences between grades were not due to the use of different tasks. To minimize working memory demands in the computerized inhibition and cognitive flexibility tasks, a reminder of response key assignment was displayed on the screen.

To avoid a priori bias about the preferable measure, we chose a broad array of tasks to measure all aspects of each executive function using different experimental paradigms. For inhibition, we chose tasks that measure resistance to distractor interference, prepotent response, and cognitive inhibition using interference paradigms (i.e., Stroop, flanker, and negative priming tasks), selective attention paradigms (i.e., receptive attention task); horse-race model (i.e., stop-signal task); and load paradigm (updating in working memory task and inhibition index). For cognitive flexibility, we used alternate (i.e., local–global), sorting (i.e., Wisconsin Card Sorting Test), and fluency (i.e., verbal fluency) paradigms that allowed us to address both perceptual and conceptual shifting. For working memory, we used both updating load paradigms (i.e., updating in working memory task and correct responses index) and span paradigms (i.e., reading and counting span tasks).

In addition, to check the discriminability of the task and ensure developmental differences, previous pilot studies with different samples at these three ages were conducted for the following reasons: to calibrate times of stimuli presentations and stimulus-onset asynchronies in response time (RT) tasks, to ensure the understandability of the task (particularly for the youngest children), to decide the number of trials, and to avoid floor and ceiling effects. The aim of these pilot studies was to detect whether the different tasks showed robust effects: consistent differences between congruent and incongruent conditions in the flanker, local–global, Stroop, and negative priming tasks; proactive interference in the updating task; and differences in task performance among the different age groups. We adjusted task parameters in response to these pilot findings, fixing them to be equal across the three age groups to avoid confounds due to subtle differences in the tasks. No task showed floor or ceiling effects. Here, we present a brief description of the task. A detailed description is presented in Appendix A in the online supplementary material.

Inhibition. We measured three types of inhibition: (a) resistance to distractor interference (receptive attention and flanker tasks); (b) prepotent response inhibition (Stroop and stop signal tasks); and (c) cognitive inhibition (negative priming and updating information in WM tasks).

Resistance to distractor interference. The *receptive attention task* is a subscale of the Cognitive Assessment System (CAS; Naglieri & Das, 1997). It measures the ability to attend to relevant stimuli selectively and ignore irrelevant stimuli. To 7-year-old children, four sheets of 200 pairs of drawings were presented. In the first condition, children were instructed to physically underline identical drawings (e.g., to underline the drawing but not the drawing XXXXX). In the second condition, children were instructed to underline drawings that belong to the same lexical category (that have the same name) (e.g., they needed to underline the drawing XXXXX but not the drawing XXXXX).

The 11- and 15-year-olds were presented with 400 pairs of letters (200 per condition). In the first condition, participants were instructed to underline physically identical pairs of letters (e.g., to underline "AA" and "aa" but not "Aa"). In the second condition, participants were instructed to underline lexically similar pairs (e.g., to underline "Aa," "AA," and "aa" but not "Tr" or "tB"). It was a paper-and-pencil task that was individually administered. The dependent variable was the base score, computed as the number of correct answers minus the number of mistakes and the time to complete the test. The test–retest reliability reported in the manual is .89. The task duration was 8 min.

The *flanker task* was an adaptation from Munro et al. (2006) and Rueda et al. (2005) in which an array of five fishes pointed to the right or left was presented. In two conditions, the participants needed to respond to the direction of the center or flanker fishes by pressing a key. In the congruent condition, the flankers were pointing in the same direction as the central fish. In the incongruent condition, the flankers pointed in the opposite direction from the central fish. The dependent variable was the RT/percentage of hits in the incongruent condition. The average duration was 20 min. Estimates of reliability for accuracy ranged from .92 (Grade 2) to .76 (adolescent sample). In the case of RT, estimates ranged from .96 (Grade 2) to .77 (adolescent sample). Some previous evidence supports the suitability of this task to measure resistance to distractor interference in children and adolescents (Huizinga et al., 2006; Rueda et al., 2005; Simonds et al., 2007; Waszak et al., 2010).

Prepotent response inhibition. The *Stroop task* was initially designed by Stroop (1935). This version was designed following MacLeod (2006). We used two types of stimuli: colored asterisks (in the neutral condition) and colored words (in the experimental condition) printed in red and blue. The participants needed to respond by naming the color of the asterisks or words vocally and pressing the intended keys. In half the experimental trials, the name of the color matched the ink color (congruent condition); in the other half, the color name did not match the ink color (incongruent condition). The dependent variable was RT/percentage of hits in incongruent conditions. The estimates of reliability ranged from .93 (Grade 2) to .78 (adolescent sample) for RT and from .96 (Grade 2) to .67 (adolescent sample) for accuracy. The average duration of the task was 10 min. Some previous evidence supported the suitability of this task to measure prepotent response inhibition in children and adolescents (Huizinga et al., 2006; Rueda et al., 2005).

In the *stop-signal task,* we used the software STOP-ITa (Verbruggen et al., 2008) to run the stop-signal paradigm (Logan & Cowan, 1984). In the no-signal conditions, the participants were instructed to respond to whether the figure was a diamond or a square by pressing a key. Under the stop-signal condition, the lines of the figures became thicker and the participants were instructed not to respond. The stop signal was presented after a variable stop-signal delay, which was adjusted following a staircase tracking procedure. The dependent variable was the stop-signal reaction time (SSRT). Reliability estimates could not be computed for this sample because the data were not recorded as failures or errors for each display. The reported reliability from simulation studies is .53 (Verbruggen et al., 2019). The average task duration was 15 min. Some previous evidence supported the suitability of this task to measure prepotent response inhibition in children and adolescents (Huizinga et al., 2006; St Clair-Thompson & Gathercole, 2006; van den Wildenberg & van der Molen, 2004).

Cognitive inhibition. For *negative priming,* we administered an adaptation of the tasks designed by Tipper (1985) and Pritchard and Neumann (2004). Along with 236 trials (20 practice and 216 experimental), drawings of objects or animals were presented in pairs of prime and probe displays. Each probe and prime display showed two overlapping shapes, green (target) and red (nontarget), and a black comparison shape presented randomly above or under the overlapping shapes. The participants were instructed to ignore the red shape and to decide whether the green shape was the same as or different from the black shape by pressing the intended keys. In the ignore condition, the stimulus to be ignored in the prime trial was the stimulus to be attended to in the probe trial. In the control condition, the stimulus to be ignored in the prime trial differed from the stimulus to be attended to in the probe trial. In half the experimental trials (control trials), the red prime distractor and the black comparison target were different; in the other half, the distractor and comparison stimuli were the same (ignored repetition trials). The dependent variable was RT/percentage of hits for the ignore trials in the probe condition. Estimates of reliability ranged from .94 (Grade 2) to .88 (adolescent sample) for RT and from .99 (Grade 2) to .86 (adolescent sample) for accuracy. The average task duration was 20 min. Some previous evidence supported the suitability of this task to measure cognitive inhibition in children and adolescents (Pritchard & Neumann, 2004; Rueda et al., 2005; Tipper, 1985).

For *updating information in WM,* an adaptation for children and adolescents of the De Beni and Palladino (2004) task was presented (Carriedo et al., 2016). Here, 24 lists of 12 words (20 experimental and 4 practice) were presented to the participants. Each list contained relevant words (to be recalled), irrelevant words (to be suppressed), and filler words. The familiarity and concreteness of the words were controlled (see Carriedo et al., 2016). The number of relevant words varied with the experimental condition: five in high-load memory conditions versus three in low-load memory conditions. The number of irrelevant words also varied with the experimental condition: two in the low-suppression condition versus five in the high-suppression condition. Words were orally presented one per second. The participants were instructed to recall the smallest animals or objects. The dependent variable was the percentage of intrusions from the same list, considered an index of suppression of information in WM, plus the percentage of previous list intrusions, considered an index of proactive interference (e.g., Carriedo et al., 2016; De Beni & Palladino, 2004; Lechuga et al., 2006; Palladino et al., 2001). The average duration of the task was 20 min. Estimates of reliability ranged from .87 (Grade 2) to .86 (adolescent sample).[2]

Cognitive flexibility. Cognitive flexibility was assessed through three tasks to address both perceptual shifting (the appearance of the stimuli to be attended to), as occurs in the local–global task, and more conceptual shifting, which included changing classification criteria, as occurs in the Wisconsin Card Sorting Test, or search criteria in long-term memory, as occurs in the fluency tasks.

Local–global task. We administered an adapted version (Montoro et al., 2011) of the Mondloch et al. (2003) task. Along with 24 practice and 48 experimental trials, the participants needed to respond to the global or local figure cued by auditory words "BIG" and "SMALL." Stimuli consisted of large squares and triangles (global figures) made up of small squares or triangles (local figures). In the congruent condition, the shape of the global and local figures matched (i.e., both were squares or triangles). In the incongruent condition, global and local shapes did not match. Local and global trials were randomly presented. The participants needed to respond to the shape of the stimulus at the level to be attended to (global or local) by pressing the intended keys. The dependent variable was RT/percentage of hits for the change conditions. The average task duration was 20 min. Estimates of reliability ranged from .90 (Grade 2) to .84 (adolescent sample) for RT and from .96 (Grade 2) to .81 (adolescent sample) for accuracy. Some previous evidence supported the suitability of this task to measure cognitive flexibility at the perceptual level in children and adolescents (Huizinga et al., 2006; Mondloch et al., 2003).

Wisconsin Card Sorting Test. We used the PEBL's Berg Card Sorting Test-64 (pBCST-64; Mueller, 2011), a computerized version of the classical Wisconsin Card Sorting Test (hereafter WCST) based on Berg (1948). The participants were instructed to classify 64 cards in terms of color, number, or shape, receiving feedback on response correctness. From this feedback, they needed to infer the classification rule. After 10 consecutive correct trials, the rule changed without warning. The dependent variable was the mean RT of the trials just after the rule change, following the same procedure as Somsen et al. (2000). Some systematic reviews have shown that the WCST is one of the most com-

[2] For details about this task and the different indexes derived from it, we direct the interested reader to the previously published article (Carriedo et al., 2016).

monly used tests to assess cognitive flexibility in children and adolescents (see Karr et al., 2018, for a review). The average task duration was 12 min. Estimates of reliability ranged from .90 (Grade 2) to .92 (adolescent sample).

Verbal fluency. We used a phonemic verbal fluency task in which the participants were instructed to generate words beginning with the letters M, P, and R for 60 s. The children were warned to avoid different verb forms, compound words, and repeat words. The order of the letters was counterbalanced. The dependent variable was the mean of the correct words produced for each letter divided by 60. Some systematic reviews have shown that the verbal fluency task is one of the most commonly used tasks to assess cognitive flexibility in children and adolescents (see Karr et al., 2018, for a review). Estimates of reliability ranged from .75 (Grade 2) to .78 (adolescent sample).

Working memory. Working memory was assessed through three tasks: two complex span tasks (reading and counting spans) and one updating task. In the reading span task, the processing task is verbal as far as it demands processing written information, whereas the counting span task demands visuospatial processing. Updating information in the WM task demands processing verbal information that was orally presented.

Reading span task for children. An adaptation for children of Daneman and Carpenter's (1980) reading span task was used (Carriedo & Rucián, 2009). A series of 48 sentences (6 practice and 42 experimental) grouped into levels of two, three, four, and five sentences (three series for each level) were presented. The participants were instructed to read each sentence aloud and remember the last word at the end of each series in the same order as presented. The task was finished when the participants failed to remember any of the series of a given level. The dependent variable was reading span, that is, the level at which the participants had correctly answered at least two of the three series. Reliability estimates cannot be computed for this sample because the data were not recorded due to failures or errors for each sentence. However, the previous hundreds of studies in which this task was used showed that reliability estimates for span scores ranged from .70 to .90 (see Conway et al., 2005, for a review).

Counting span task. The counting span task has the same structure as Daneman and Carpenter's (1980) reading span task, but in this task the participants needed to process visual information (counting and pointing out geometric figures). The current version was an adaptation of Case et al. (1982). A series of 48 displays (6 practice and 42 experimental) grouped into levels of two, three, four, and five displays (three series for each level) were presented. Each display was formed of 18 red and blue squares. The participants were instructed to count the blue squares of each display and remember the number of squares counted in the same order in which they were presented. The task was finished when the participants failed to remember any of the series of a given level. Reliability estimates could not be computed for this sample because the data were not recorded due to failures or errors for each display. However, the previous hundreds of studies in which this task was used showed that reliability estimates for span scores ranged from .70 to .90 (see Conway et al., 2005, for a review).

Updating information in the WM task. We used the adaptation of De Beni and Palladino (2004; see also Palladino et al., 2001) adapted for children and adolescents depicted above. As an index of updating, the proportion of correct responses was used.

Fluid intelligence

The total raw score obtained in Raven's Progressive Matrices SPM was used as a measure of fluid intelligence. This scale comprises 60 items organized into five sets (A–E) of 12 items each. For each item, the children needed to complete a series of complex spatial figures using analogical reasoning. In addition, the test provided centile scores and grades, which were used in the initial selection of the sample. The reliability estimates were .95 and .93, respectively, for the Grade 2 and adolescent samples.

Processing speed

We used, as an estimator of processing speed, the RT/percentage of hits in the neutral condition of the Stroop task (Appendix A). The reliability estimates for RT and accuracy were .93 and .78 for the Grade 2 sample and .96 and .67 for the adolescent sample, respectively.

Procedure

Before starting the research, the schools and families were informed of the study's objectives and the activities in which the students would participate. The students' families gave their written consent before testing began. The university's ethics committee approved all methods included in this study.

The assessment was carried out in four sessions at schools during the second semester. First, the executive tasks were assessed individually in a quiet space. Except for the receptive attention task, the stimuli were administered by a computer. Stimulus randomization and timing were controlled using E-Prime software (Version 2.0; Schneider et al., 2002). Fluid intelligence and mathematical tasks were assessed collectively in the classroom following the instructions from the manuals. Flankers, reading span, and receptive attention were administered in the first session. The second session included local–global, WCST, counting, and go/no-go tasks. The third session included Stroop, stop-signal, and updating tasks. The fourth session included BADyG and Raven tasks. Counterbalancing was performed following the Latin square procedure per session.

Analyses

Before the main analyses, the distribution of variables and outliers was examined according to the procedure established by Friedman et al. (2009). In each of the school grades, outliers that deviated more than 3 standard deviations from the mean were trimmed to 3. This affected 1.6% of the RTs for Grades 2 and 10 and 1.47% of the scores for Grade 6. All scores were then transformed into *z* scores.

Several sets of analyses were conducted for the three grades. First, to identify the best combination of executive predictors of mathematical achievement, a Bayesian factor-analytic approach (Rouder & Morey, 2015) was used through the JASP software package (Version 0.16.3; JASP Team, 2020). At each school grade, a structured sequence of four Bayesian regression analyses was conducted to gather step-by-step information on the best predictors linked to each of the EFs (inhibition, cognitive flexibility, and WM) until a final combined analysis that encompassed the best set of executive variables providing concrete information on the relative importance of each of the variables was reached.

Second, in each school grade, various binary logistic regression analyses were carried out using the Jamovi statistical platform (Version 2.3.2; The Jamovi Project, 2021) to analyze the predictive ability of the executive variables that were relevant in the previous Bayesian models to classify students in accordance with their previous mathematical achievement (mathematical difficulties vs. typical achievement).

Third, processing speed and intelligence were included to test the predictive capacity of the executive models concerning broader models that included these other two cognitive variables. Finally, these broader models were reduced step by step to obtain the most parsimonious combined model for each grade.

Results

Tables B1 (Grade 2), B2 (Grade 6), and B3 (Grade 10) in Appendix B in the online supplementary material present the descriptive statistics and correlations among the study variables for the age groups. Table B4 shows the mean scores for the achievement groups.

Bayesian regressions

At each school grade, the Bayes factor (BF_{10}) was calculated to compare the models of inhibition (M_{IN}), cognitive flexibility (M_{CF}), and WM (M_{WM}) and the final combined model (M_{EF}) with the null model (BF_{NM}), which did not include predictors, and with the top model (BF_{TM}), which presented the highest Bayes factor in each case. Comparison of the models of interest with the top model made it possible to determine the suitability of the combination of variables included in the top model and to quantify the importance of each of them. Bayes factors below .33 provide substantial evidence in favor of differences between the models, those below .10 indicate strong evidence, those from .03 to .01 indicate very strong evidence, and those below .01 are decisively in favor of the model or the inclusion of the variable (Jeffreys, 1961).

We present only a summary of the top Bayesian models for the prediction of mathematical achievement in Tables 1 (Grade 2), 2 (Grade 6), and 3 (Grade 10).

Grade 2

The top inhibition model (M_{IN_TOP}) included only two of the six possible variables (Table 1): receptive attention and negative priming (BF_{NM} = 743.23, R^2 = .14). Comparison of the model with alternative models (M_{IN2}–M_{IN3}) provided very strong evidence for the retention of negative priming (M_{IN3}) as a predictor of mathematical achievement (BF_{TM} = .01).

The top model of cognitive flexibility (M_{FC_TOP}) included the effects of two variables: local–global and verbal fluency (BF_{NM} = 30215.13, R^2 = .19). The exclusion of both in the following models (M_{CF2} and M_{CF3}) provided very strong evidence for retaining local–global as a predictor (BF_{TM} = .02) and decisive evidence for verbal fluency (BF_{TM} < .01).

The top model of WM ($_{MWM\_TOP}$) included only counting span as a predictor (BF_{NM} = 710.39, R^2 = .06).

The top model of executive functioning (M_{EF_TOP}) included the effects of three variables: negative priming, local–global, and verbal fluency (BF_{NM} = 1.26 × 10^6, R^2 = .25). Comparison of the top model with alternative models that sequentially excluded the variables (M_{EF2}–M_{EF4}) provided substantial evidence for retaining local–global as a predictor (BF_{TM} = .21), very strong evidence for negative priming (BF_{TM} = .02), and decisive evidence for verbal fluency (BF_{TM} < .01).

Grade 6

The top inhibition model (M_{IN_TOP}) included five of the six variables (Table 2): flankers, receptive attention, stop signal, negative priming, and intrusions (BF_{NM} = 8.81 × 10^7, R^2 = .33). Comparison of the model with alternative models that sequentially excluded the variables (M_{IN2}–M_{IN6}) provided strong evidence for retaining flankers as a predictor of mathematical achievement (BF_{TM} = .05) and decisive evidence for receptive attention (BF_{TM} < .01).

The top model of cognitive flexibility (M_{FC_TOP}) included the effects of two of the three variables: local–global and verbal fluency (BF_{NM} = 3.48 × 10^7, R^2 = .27). Comparison of the model with alternative models (M_{CF2} and M_{CF3}) that excluded the two variables provided very strong evidence for retaining local–global as a predictor (BF_{TM} = .09) and decisive evidence for verbal fluency (BF_{TM} < .01).

The top model of WM (M_{WM_TOP}) included only counting span as a predictor (BF_{NM} = 4.53 × 10^7, R^2 = .25).

Finally, the top model of executive functioning (M_{EF_TOP}) included only the effects of receptive attention, local–global, and counting span (BF_{NM} = 4.09 × 10^{13}, R^2 = .42). Comparison of the top model with alternative models that sequentially excluded the variables (M_{EF2}–M_{EF4}) provided very strong evidence for retaining receptive attention as a predictor (BF_{TM} < .03) and decisive evidence for local–global and counting span (BF_{TM} < .01).

Grade 10

The top inhibition model (M_{IN_TOP}) included the effects of three of the six variables (Table 3): flankers, receptive attention, and stop signal (BF_{NM} = 140815.98, R^2 = .27). Comparison of the model with alternative models that sequentially excluded the variables (M_{IN2}–M_{IN4}) provided substantial evidence for retention as a predictor of flankers (BF_{TM} = .30) and receptive attention (BF_{TM} = .21) and very strong evidence for stop signal (BF_{TM} < .03).

The top model of cognitive flexibility (M_{CF_TOP}) included two of the three variables: local–global and WCST (BF_{NM} = 34061, R^2 = .10). Comparison with the following models (M_{CF2} and M_{CF3}) provided substantial evidence only for the retention of local–global (BF_{TM} = .11).

The top model of WM (M_{WM_TOP}) included the variables reading span and updating as predictor variables (BF_{NM} = 74641, R^2 = .11). Comparison with the following models (M_{WM2} and M_{WM3}) provided substantial evidence for reading span retention as a predictor of mathematical achievement (BF_{TM} = .16).

Table 1
Bayes factor analysis of predictors of mathematical achievement (Grade 2)

Model	Variables		Bayes factor		R^2
	Included	Excluded	BF_{NM}	BF_{TM}	
Inhibition					
M_{IN_TOP}	Receptive attention + negative priming	–	743.23	1	.14
M_{IN2}	Negative priming	Receptive attention	313.59	.42	.11
M_{IN3}	Receptive attention	Negative priming	10.20	**.01**	.06
Cognitive flexibility					
M_{CF_TOP}	Local–global + verbal fluency	–	30215.13	1	.19
M_{CF2}	Verbal fluency	Local–global	620.15	**.02**	.12
M_{CF3}	Local–global	Verbal fluency	176.94	**<.01**	.10
Working memory					
M_{WM_TOP}	Counting span	–	10.39	1	.06
Executive functioning					
M_{EF_TOP}	Negative priming + local–global + verbal fluency	–	1.26×10^6	1	.25
M_{EF2}	Negative priming + verbal fluency	Local–global	263074.16	**.21**	.22
M_{EF3}	Local–global + verbal fluency	Negative priming	30215.13	**.02**	.19
M_{EF4}	Negative priming + local–global	Verbal fluency	4625.11	**<.01**	.17

Note. M_{IN}, inhibition model; M_{CF}, cognitive flexibility model; M_{WM}, working memory model; M_{EF}, executive functioning model; BF_{NM}, Bayes factor between particular models and the null model; BF_{TM}, Bayes factor between particular models and the top model. Bold values indicate significance.

Table 2
Bayes factor analysis of predictors of mathematical achievement (Grade 6)

Model	Variables		Bayes factor		R^2
	Included	Excluded	BF_{NM}	BF_{TM}	
Inhibition					
M_{IN_TOP}	Flankers + receptive attention + stop signal + negative priming + intrusions	–	8.81×10^7	1	.33
M_{IN2}	Flankers + receptive attention + stop signal + intrusions	Negative priming	7.31×10^7	.83	.31
M_{IN3}	Flankers + receptive attention + stop signal + negative priming	Intrusions	4.57×10^7	.52	.31
M_{IN4}	Flankers + receptive attention + negative priming + intrusions	Stop signal	3.08×10^7	.35	.30
M_{IN5}	Receptive attention + stop signal + negative priming + intrusions	Flankers	4.79×10^6	**.05**	.28
M_{IN6}	Flankers + stop signal + negative priming + intrusions	Receptive attention	249375.547	**<.01**	.25
Cognitive flexibility					
M_{CF_TOP}	Local–global + verbal fluency	–	3.48×10^7	1	.27
M_{CF2}	Local–global	Verbal fluency	3.20×10^6	**.09**	.22
M_{CF3}	Verbal fluency	Local–global	2.01×10^6	$\mathbf{2.25 \times 10^{-6}}$	.09
Working memory					
M_{WM_TOP}	Counting span	–	4.53×10^7	1	.25
Executive functioning					
M_{EF_TOP}	Receptive attention + local–global + counting span	–	4.09×10^{13}	1	.42
M_{EF2}	Local–global + counting span	Receptive attention	1.17×10^{12}	**<.03**	.37
M_{EF3}	Receptive attention + counting span	Local–global	9.70×10^{10}	**<.01**	.35
M_{EF4}	Receptive attention + local–global	Counting span	4.86×10^8	$\mathbf{1.19 \times 10^{-5}}$	.30

Note. M_{IN}, inhibition model; M_{CF}, cognitive flexibility model; M_{WM}, working memory model; M_{EF}, executive functioning model; BF_{NM}, Bayes factor between particular models and the null model; BF_{TM}, Bayes factor between particular models and the top model.

Table 3
Bayes factor analysis of predictors of mathematical achievement (Grade 10)

Model	Variables		Bayes factor		R^2
	Included	Excluded	BF_{NM}	BF_{TM}	
Inhibition					
M_{IN_TOP}	Flankers + receptive attention + stop signal	–	140815.98	1	.27
M_{IN2}	Receptive attention + stop signal	Flankers	42804.35	**.30**	.19
M_{IN3}	Flankers + stop signal	Receptive attention	29358.93	**.21**	.19
M_{IN4}	Flankers + receptive attention	Stop signal	3747.33	**<.03**	.16
Cognitive flexibility					
M_{CF_TOP}	Local–global + WCST	–	34.061	1	.10
M_{CF2}	Local–global	WCST	20.522	.60	.07
M_{CF3}	WCST	Local–global	3.616	**.11**	.04
Working memory					
M_{WM_TOP}	Reading span + updating	–	74.641	1	.11
M_{WM2}	Reading span	Updating	29.678	.40	.08
M_{WM3}	Updating	Reading span	11.704	**.16**	.06
Executive functioning					
M_{EF_TOP}	Flankers + receptive attention + stop signal + reading span	–	775725.39	1	.26
M_{EF2}	Receptive attention + stop signal + reading span	Flankers	370662.00	.48	.24
M_{EF3}	Flankers + stop signal + reading span	Receptive attention	150573.09	**.19**	.23
M_{EF4}	Flankers + receptive attention + stop signal	Reading span	140815.98	**.18**	.23
M_{EF5}	Flankers + receptive attention + reading span	Stop signal	38950.28	**.05**	.21

Note. M_{IN}, inhibition model; M_{CF}, cognitive flexibility model; M_{WM}, working memory model; M_{EF}, executive functioning model; BF_{NM}, Bayes factor between particular models and the null model; BF_{TM}, Bayes factor between particular models and the top model; WCST, Wisconsin Card Sorting Test.

The top model of executive functioning (M_{EF_TOP}) included the effects of four of the five variables selected in the previous analyses: flankers, receptive attention, stop signal, and reading span (BF_{NM} = 775725.39, R^2 = .26). Comparison of the top model with the following models (M_{EF2}–M_{EF5}) provided substantial evidence for retaining receptive attention (BF_{TM} = .19) and reading span (BF_{TM} = .18) as predictors of mathematical achievement and very strong evidence for stop signal (BF_{TM} = .05).

Logistic regressions

Tables 4 (Grade 2), 5 (Grade 6), and 6 (Grade 10) present only three models for each grade: (a) Model 1, which included the executive variables resulting from the top model of EF as predictors; (b) Model 2, with processing speed and intelligence added; and (c) the final more parsimonious model. In all analyses, the typical achievement group was used as the reference group in the contrasts. The inclusion of interactions was not considered given that previous analyses did not show significant interaction effects. When selecting the best model, sensitivity (% participants with correctly classified mathematical difficulties), specificity (% participants with correctly classified typical achievement), and the receiver-operating characteristic (ROC) curve were analyzed as a measure of model efficiency. Values from .80 to .90 indicate excellent discrimination (Hosmer et al., 2013).

Grade 2

The first model (Table 4), which included negative priming (cognitive inhibition), local–global (cognitive flexibility), and verbal fluency (cognitive flexibility), was statistically significant, R^2_N = .25, $\chi^2(3)$ = 20.6, p < .001, and correctly classified 78% of the mathematical difficulties group and 79% of the typical achievement group. The area under the curve (AUC) of .82 was excellent. Negative priming (p = .029) and verbal fluency (p = .003) were the variables that best discriminated between groups. A 1-unit increase in the negative priming score increased the odds of presenting mathematical difficulties 1.86 times, odds ratio (OR) = 1.86, 95% confidence interval (CI) [1.06, 3.25]. On the other hand, a 1-unit

Table 4
Results of binary logistic regressions (Grade 2)

Variable	Model 1					Model 2					Final model				
	Estimate	*SE*	*Z*	*p*	OR	Estimate	*SE*	*Z*	*p*	OR	Estimate	*SE*	*Z*	*p*	OR
Intercept	−2.37	0.36	−6.61	<.001	0.09	−2.67	0.44	−6.11	<.001	0.07	−2.54	0.40	−6.41	<.001	0.08
Negative priming	0.62	0.28	2.18	.029	1.86	0.45	0.32	1.42	.155	1.58					
Local–global	0.36	0.25	1.41	.157	1.43	0.25	0.26	0.93	.352	1.28					
Verbal fluency	−0.99	0.33	−3.00	.003	0.37	−0.97	0.38	−2.58	.010	0.38	−0.90	0.36	−2.51	.012	0.41
Processing speed						0.64	0.27	2.36	.018	1.90	0.79	0.25	3.12	.002	2.19
Intelligence						−0.72	0.38	−1.87	.062	0.49	−0.76	0.37	−2.08	.037	0.47
R^2_N	.25					.37					.33				
Sensitivity	.78					.77					.78				
Specificity	.79					.80					.76				
AUC	.82					.88					.85				

Note. Z, *z*-statistic; OR, odds ratio; R^2_N, Nagelkerke's R^2; AUC, area under the curve.

increase in verbal fluency score reduced the odds of presenting mathematical difficulties 2.70 times, OR = 0.37, 95% CI [0.19, 0.71].

The introduction in Model 2 of processing speed and intelligence significantly increased variance, $\Delta R^2_N = .12$, $\chi^2(2) = 10.6$, $p = .005$, although sensitivity, specificity, and AUC remained in similar ranges. In this model, only verbal fluency ($p = .010$) and processing speed ($p = .018$) contributed significantly to discriminating between groups.

The successive elimination of variables that were not significant in the previous analysis resulted in different models. The most parsimonious model included verbal fluency, processing speed, and intelligence, $R^2_N = .33$, $\chi^2(3) = 27.4$, $p < .001$, and correctly classified 78% of the mathematical difficulties group and 76% of the typical achievement group, percentages similar to those of the previous models. The AUC also remained in the same range (AUC = .82). Processing speed was the main risk factor, OR = 2.19, 95% CI [1.34, 3.59], with a 1-unit increase in processing speed increasing the odds of presenting mathematical difficulties 2.19 times. On the other hand, intelligence, OR = 0.46, 95% CI [0.19, 0.96], and verbal fluency, OR = 0.41, 95% CI [0.20, 0.82], acted as protective factors given that an increase of 1 unit in their score reduced the odds of presenting mathematical difficulties by 2.13 and 2.43, respectively.

Grade 6

The executive model (Table 5) that included receptive attention (resistance to distractor interference), local–global attention (cognitive flexibility), and counting (WM) was statistically significant, $R^2_N = .36$, $\chi^2(3) = 41.4$, $p < .001$, and correctly classified 71% of the mathematical difficulties group and 75% of the typical achievement group. The AUC of .82 was excellent. Local–global ($p < .001$) and counting span ($p = .003$) were the variables that discriminated between groups. A 1-unit increase in the local–global score increased the odds of presenting mathematical difficulties 2.45 times, OR = 2.45, 95% CI [1.51, 3.99]. On the other hand, a 1-unit increase in the counting span score reduced the odds of presenting mathematical difficulties 1.75 times, OR = 0.45, 95% CI [0.28, 0.74].

The inclusion of cognitive variables in Model 2 significantly increased variance, $\Delta R^2_N = .05$, $\chi^2(2) = 7.65$, $p = .022$, although sensitivity, specificity, and AUC remained in similar ranges. In this model, local–global ($p = .0002$), counting span ($p = .041$), and intelligence ($p = .010$) contributed significantly to discriminating between groups.

The final model included these three variables, $R^2_N = .40$, $\chi^2(3) = 47.9$, $p < .001$, and correctly classified 76% of the mathematical difficulties group and 75% of the typical achievement group, percentages similar to those of the previous models. The AUC also remained in the same range (AUC = .84). Local–global was the main risk factor, with a 1-unit increase in the local–global score increasing the odds of presenting mathematical difficulties 2.52 times, OR = 2.52, 95% CI [1.54, 4.12]. On the other hand, a 1-unit increase in intelligence score, OR = 0.42, 95% CI [0.23, 0.76], and counting span, OR = 0.55, 95% CI [0.33, 0.94], reduced the odds of experiencing mathematical difficulties by 2.38 and 1.83, respectively, thereby acting as protective factors.

Grade 10

The executive model (Table 6) that included receptive attention (resistance to distractor interference), stop signal (prepotent response inhibition), and reading span (WM) was statistically significant, $R^2_N = .33$, $\chi^2(3) = 33.1$, $p < .001$, and correctly classified 79% of the mathematical difficulties group and 78% of the typical achievement group. The AUC of .82 placed the discrimination ability as excellent. Stop signal ($p < .001$) and reading span ($p = .003$) were the variables that discriminated between the groups. A 1-unit increase in the stop-signal score increased the odds of presenting mathematical difficulties 2.85 times, OR = 2.85, 95% CI [1.78, 4.58]. On the other hand, a 1-unit increase in reading span score reduced the odds of presenting mathematical difficulties 1.78 times, OR = 0.56, 95% CI [0.35, 0.91].

The inclusion of processing speed and intelligence in Model 2 significantly increased variance, $\Delta R^2_N = .03$, $\chi^2(2) = 6.48$, $p = .039$, although sensitivity, specificity, and AUC remained at similar percentages. In this model, stop signal ($p < .001$) and intelligence ($p = .017$) contributed significantly to the prediction of mathematical achievement.

Successive elimination of the variables resulted in a final model that included only these two variables, $R^2_N = .36$, $\chi^2(3) = 36.6$, $p < .001$. This model correctly classified 78% of the mathematical difficulties group and 77% of the typical achievement group, percentages similar to those of the previous models.

Table 5
Results of binary logistic regressions (Grade 6)

Variable	Model 1					Model 2					Final model				
	Estimate	*SE*	*Z*	*p*	OR	Estimate	*SE*	*Z*	*p*	OR	Estimate	*SE*	*Z*	*p*	OR
Intercept	−1.16	0.24	−4.88	<.001	0.31	−1.15	0.25	−4.64	<.001	0.32	−1.11	0.24	−4.63	<.001	0.33
Receptive attention	−0.36	0.26	−1.42	.157	0.70	−0.22	0.27	−0.83	.407	0.80					
Local–global	0.90	0.25	3.61	<.001	2.45	0.83	0.27	3.09	.002	2.29	0.93	0.25	3.70	<.001	2.52
Counting span	−0.79	0.25	−3.13	.002	0.45	−0.56	0.27	−2.05	.041	0.57	−0.59	0.27	−2.18	.029	0.55
Processing speed						0.18	0.25	0.71	.475	1.19					
Intelligence						−0.81	0.31	−2.58	.010	0.45	−0.86	0.30	−2.83	.005	0.42
R^2_N	.36					.41					.40				
Sensitivity	.71					.74					.76				
Specificity	.75					.75					.75				
AUC	.82					.85					.84				

Note. Z, *z*-statistic; OR, odds ratio; R^2_N, Nagelkerke's R^2; AUC, area under the curve.

The AUC also remained in the same range (AUC = .84). Stop signal was the main risk factor, with a 1-unit increase in its score increasing the odds of presenting mathematical difficulties 3.09 times, OR = 3.09, 95% CI [1.91, 5.03]. On the other hand, intelligence acted as a protective factor, with a 1-unit increase in intelligence score reducing the odds of presenting mathematical difficulties 2.70 times, OR = 0.37, 95% CI [0.19, 0.71].

Discussion

This study analyzed the contributions of different measures of EFs to academic achievement in Grades 2, 6, and 10. Overall, the results revealed differential contributions of EFs across the three selected grades, the implications of which are discussed in terms of the two initial goals.

Executive predictors of academic achievement in mathematics

Based on a Bayesian factor-analytic approach, our first goal was to analyze the extent to which individual differences in executive measures of inhibition, cognitive flexibility, and WM were associated with individual differences in mathematical achievement in Grades 2, 6, and 10. Overall, the results of the study were consistent with our initial expectations that EF measures contributed differentially to academic achievement in Grades 2, 6, and 10. Moreover, Bayesian analyses provided novel data on different combinations of predictors at each grade. For Grade 2, the final combination of variables included a measure of cognitive inhibition (negative priming) and two measures of cognitive flexibility (local–global and verbal fluency). The Grade 6 model included as predictors a measure of resistance to distractor interference (receptive attention), a measure of cognitive flexibility (local–global), and a measure of WM (counting span). Finally, in Grade 10 the model included two measures of inhibition related to resistance to distractor interference and prepotent response inhibition (receptive attention and stop signal) and one measure of WM (reading span).

These results are not congruent with those obtained by Lee and Bull (2016) and Cragg et al. (2017), who reduced the contribution of EFs to mathematical achievement to WM. There may be several reasons for this discrepancy. One possibility may be the different mathematical tasks used in the three studies. In the Cragg et al. (2017) study, greater emphasis was placed on mathematical reasoning ability relative to the more curricular tasks used in the Lee and Bull (2016) study and this study. It is also possible that the differences in the number and nature of the selected executive tasks among the three studies contribute to the observed discrepancies. In this study, a larger number of tasks were included, and measures of cognitive inhibition and verbal fluency were included. Another potential explanation for the discrepancies could be the analytical approach taken when considering executive variables

Table 6
Results of binary logistic regressions (Grade 10)

Variable	Model 1					Model 2					Final model				
	Estimate	*SE*	*Z*	*p*	OR	Estimate	*SE*	*Z*	*p*	OR	Estimate	*SE*	*Z*	*p*	OR
Intercept	−1.88	0.30	−6.31	<.001	0.15	−1.92	0.31	−6.08	<.001	0.15	−1.81	0.29	−6.24	<.001	0.16
Receptive attention	−0.17	0.25	−0.68	.495	0.84	−0.11	0.27	−0.42	.675	0.89					
Stop signal	1.05	0.24	4.34	<.001	2.85	1.14	0.26	4.36	<.001	3.14	1.13	0.25	4.57	<.001	3.10
Reading span	−0.57	0.24	−2.34	.019	0.56	−0.35	0.26	−1.31	.189	0.71					
Processing speed						−0.27	0.28	−0.97	.333	0.76					
Intelligence						−0.89	0.38	−2.38	.017	0.41	−1.00	0.34	−2.94	.003	0.37
R^2_N						.39					.36				
Sensitivity	.79					.79					.78				
Specificity	.78					.79					.77				
AUC	.82					.85					.84				

Note. Z, *z*-statistic; OR, odds ratio; R^2_N, Nagelkerke's R^2; AUC, area under the curve.

together. In contrast to the preliminary correlational study with factor variables by Lee and Bull (2016) and the joint regression analysis by Cragg et al. (2017), this study employed a Bayesian individual differences approach for each age group to identify the optimal combination of executive measures that explain individual differences in curricular achievement in mathematics and to quantify the relative importance of each. Further research is needed to fully understand the sources of these discrepancies.

Globally, the results place us in a more dynamic context where the relationship between EF measures and mathematical achievement seems to vary throughout the school grades analyzed, as we explain below, according to the development of EFs and the demands of mathematical tasks (Bull & Lee, 2014).

Specifically, in line with our first prediction, Bayesian analyses found different inhibition dimensions (cognitive and resistance to distractor interference) to be stable predictors of mathematical achievement across the three grades. Bayes factors showed a very strong contribution of receptive attention, linked to resistance to distractor interference in Grade 6, and the stop-signal task, linked to prepotent response inhibition in Grade 10. Receptive attention also contributed substantially to academic achievement in Grade 10. These results are congruent with findings from other studies that have associated response inhibition with mathematical achievement in late primary school grades (e.g., Gerst et al., 2015; Lubin et al., 2016) and occasionally in adolescent samples (Latzman et al., 2010). Importantly, Bayesian analysis showed that the effect of response distractor inhibition on explaining math achievement is not isolated. It is combined with measures of cognitive flexibility in Grades 2 and 6 and with measures of WM in Grades 6 and 10. In addition, negative priming, a paradigm traditionally associated with cognitive inhibition, emerged as a very strong predictor of individual differences in mathematical achievement in Grade 2, an effect not previously reported to our knowledge. Thus, this paradigm, widely used in samples of participants with cognitive disorders, may provide important information for studying the resolution of complex mathematical tasks (e.g., those that include multiple steps). These tasks require students to resist proactive interference of information that may have been relevant in the initial steps of the task but that has ceased to be so for the ongoing step.

Another interesting result, in line with our prediction, was associated with the contribution of cognitive flexibility to mathematical achievement throughout primary school (Nunes de Santana et al., 2022; Yeniad et al., 2013). Bayes factors pointed to the decisive contribution of verbal fluency in Grade 2 and the substantial contribution of local–global fluency in Grade 2 and decisive contribution in Grade 6. However, no predictive relationships with mathematical achievement were found in Grade 10, in line with previous findings (Boschloo et al., 2014; Cragg et al., 2017). Taken together, the results provide two main contributions. First, we found predictive relationships between cognitive flexibility and mathematical achievement that go beyond the relationship with arithmetic skills previously established in the literature (see Nunes de Santana et al., 2022, and Yeniad et al., 2013, for reviews). Moreover, contrary to what has been reported by some previous studies, these relationships were maintained in the presence of the other executive variables in Grades 2 and 6 (Cragg et al., 2017; Lee & Bull, 2016). These results could be explained in terms of changes in the demands of mathematical and executive tasks linked to age and schooling (Friso-van den Bos et al., 2013). The mathematical tasks used in Grades 2 and 6 seem to imply strong cognitive demands for students associated with the increased difficulty of arithmetic problems that would not yet be procedurally consolidated (Grade 2) and the incorporation of complex series and problems involving mathematical concepts such as proportions, decimals, and algebra (Grade 6). The lack of predictive relationships and the decrease in associations in Grade 10 (Boschloo et al., 2014; Cragg et al., 2017) could be associated with the fact that the math and executive tasks were not sufficiently demanding for these students. In the case of math tasks, although those used for 10th graders involved more complex calculations than those used for 6th graders, they are similar from a procedural point of view. Cognitive flexibility tasks, such as verbal fluency, could be less demanding for 15-year-old students. Finally, in line with the literature, we observed a significant contribution of WM to mathematical achievement (Friso-van den Bos et al., 2013), although it was less stable than expected according to previous studies (Cragg et al., 2017; Lee & Bull, 2016). Bayes factors established a very strong contribution of the counting span measure at Grade 6 and a substantial contribution of the reading span measure at Grade 10. In Grade 2, the results of the Bayesian analyses reflect an anecdotal contribution when the models included measures of inhibition (negative priming) and cognitive flexibility (verbal fluency and local–global). This lack of effect of WM in Grade 2 could be due to the nature of the arithmetic task used in this grade or the higher representation of students with typical mathematical achievement (87.33% of the sample) who tend to rely less on WM capacity when solving mathematical tasks and use more efficient WM strategies (e.g., better number encoding strategies, better verbal or visuospatial representations) (Friso-van den Bos et al., 2013).

Cognitive functioning and mathematical achievement groups

Our second goal was to analyze the ability of EFs to combine with each other and with other related constructs, such as fluid intelligence and processing speed, to classify students according to their academic achievement in mathematics. Overall, our results showed that the executive models derived from the Bayesian analyses had a similar ability to classify students with mathematical difficulties and their peers with typical mathematical achievement to the broader cognitive models that include fluid intelligence and processing speed, confirming our prediction. Therefore, consistent with the results of the meta-analysis by Peng et al. (2018), EFs appear to be positioned as consistent differentiators of achievement groups throughout the school grades analyzed. From the educational point of view, the study's main contribution is related to the differential roles of the different EF measures, fluid intelligence, and processing speed throughout schooling when their effects are considered together.

Specifically, the combination of measures that best discriminated between groups in Grade 2 included processing speed, cognitive flexibility (verbal fluency), and fluid intelligence. In this school grade, the group selection task was arithmetic, and in line with our initial prediction, processing speed was the main risk factor in predicting mathematical difficulties. This result is consistent with previous studies that placed this deficit as intrinsic to arithmetic difficulty (e.g., Cirino et al., 2015; Vukovic & Siegel, 2010). According to our predictions and previous research, IQ and verbal fluency act as protective factors given that a higher IQ and better verbal fluency reduce the probability of presenting mathematical difficulties. The finding that verbal fluency (cognitive flexibility) was a protective factor was consistent with our prediction and previous findings (McLean & Hitch, 1999; Szucs et al., 2013). In the same vein, the finding that intelligence is a protective factor is consistent with Geary's (2013) account of the facilitating role of intelligence in comprehending abstract mathematical information.

In Grade 6, the most parsimonious combination of predictors comprised local–global (cognitive flexibility), counting span (WM), and fluid intelligence measures. In this grade, the main risk factor for mathematical difficulties was performance on the local–global task. This finding could reflect the difference between the attentional shifting demanded by mathematical tasks and the participants' actual cognitive flexibility developmental level (see Huizinga et al., 2006). This discrepancy is especially significant in mathematical tasks involving a number of different structures, operations, and procedures that require multiple steps and changing strategies during their resolution (Archambeau & Gevers, 2018).

Moreover, consistent with the literature and our hypothesis, in Grade 6 the counting span task, linked to WM verbal ability, contributed to discriminating between the mathematical difficulties group members and their peers (e.g., Cirino et al., 2015; Peng et al., 2018). Counting span and fluid intelligence emerged as protective factors for the probability of presenting mathematical difficulties, which is congruent with previous literature that pointed out the facilitative role of intelligence (Geary, 2013) and WM (e.g., Bull & Lee., 2014; Friso-van den Bos et al., 2013) for math achievement, especially during the resolution of complex mathematical tasks that demand different steps, procedures, and operations.

Finally, in Grade 10, the combination of predictors that best discriminated between the achievement groups included stop signal, a measure of prepotent response inhibition (a risk factor for mathematical difficulties), and fluid intelligence (as a protective factor for mathematical difficulties). It is worth mentioning that worse performance in the stop-signal task increased the probability of mathematical difficulties by more than three times, which underlines the importance of inhibiting the prepotent response of applying an overlearned procedure when facing arithmetical problem solving. These results partially confirm our initial prediction and extend the results of previous studies in which prepotent response inhibition assessed with this same task differentiated students with mathematical difficulties from controls (e.g., Szucs et al., 2013; Willcutt et al., 2013) and contributed to predicting their difficulties (Willcutt et al., 2013).

In any case, according to our predictions, we would expect inhibition and WM to contribute to a greater ability to discriminate groups. The results of the univariate analyses were along these lines, differentiating the groups globally across the three school grades. However, the demanding analytical approach used first in the selection of the executive variables and then in the search for the most parsimonious combined models places us in a more selective and dynamic scenario where the roles of EFs, processing speed, and intelligence seem to vary according to their interrelationships and the demands of the mathematical tasks.

Limitations and future research lines

This study should be viewed in light of the measures used and the correlational nature of the data. This design has allowed the establishment of some predictors that acted as protective or risk factors for mathematical achievement. However, for a deeper comprehension of the predictive value of the different executive measures and how they emerge through Grades 2, 6, and 10, it would be necessary to replicate these results using a longitudinal design that allows us to compare the performance of the same participants across grades. Moreover, it would require assessing mathematical performance through specific mathematical tasks or the design of ad hoc experimental tasks that help us to establish hierarchical relationships between each executive measure and the different domain-specific skills, such as mental and written arithmetic, problem solving, and algebra. The use of this kind of task would provide more accurate data at each processing level and essential information about the strategies used by the students and then the most efficient strategies to work in the classroom by the teachers.

Furthermore, it is important to be cautious given that it is not clear that students at the same age strategically cope with executive tasks in similar ways (Huizinga et al., 2006). In addition, under discussion is whether executive tasks demand different EFs during development (Gilmore & Cragg, 2018) and to what extent variables such as impulsivity may contribute to explaining their performance (Skippen et al., 2019). In this sense, although an important effort has been made in the initial screening of participants and the design of the assessment, we cannot rule out that the trend shown by the stop-signal scores of adolescent students with mathematical difficulties may be associated with problems in impulsivity control. In this case, it would be important to consider valid objective and subjective measures in future research that would allow us to analyze the effect of impulsivity on stop-signal performance and, consequently, on mathematical achievement.

On the other hand, given that the literature has considered the influence of visuospatial skills, oral language, and domain-specific skills on mathematical achievement (see Peng et al., 2018, for a review), it would be interesting for future studies to analyze how these skills interact with EF skills in predicting mathematical achievement and difficulty.

Finally, one of the main problems in evaluating EF is the task impurity problem; that is, the tasks used are not specific to executive processes. One of the methods proposed to address the impurity problem is confirmatory factor analysis (Miyake et al., 2000). However, this method does not eliminate the possibility that the factor structure also reflects nonexecutive processes (van der Ven, 2011; van der Ven et al., 2012), and frequently very similar tasks are included as indicators of the same latent variable. Another approach is to use control tasks and calculate difference scores between control (congruent) conditions and experimental (incongruent) conditions to use as dependent variables. However, the use of difference scores is often criticized in the literature due to their problems of over-additivity (Faust et al., 1999) and lack of reliability (e.g., Friedman & Miyake, 2004; Hedge et al., 2018; Kane et al., 2016; Paap & Sawi, 2016; Redick et al., 2016; Rey-Mermet et al., 2018; Unsworth & McMillan, 2014; Unsworth et al. 2020). Therefore, there is no perfect approach to addressing the task impurity problem.

To minimize this problem in the current study, we selected 12 tasks validated by confirmatory factor analysis in previous studies through a careful literature review. In addition, the tasks chosen for each hypothetical construct used different stimulus modalities (verbal and visuospatial), were representative of different experimental paradigms, and did not include numerical or mathematical content to avoid overlaps with the criterion variable and the influence of domain-specific content. Moreover, we also controlled WM demands, processing speed, and fluid intelligence, which are other sources of impurity of EF tasks.

We acknowledge that this set of strategies only partially addresses the task impurity problem; however, we believe that our study has an advantage over the majority of studies in the literature on executive functioning and academic achievement because most of them used only a single task to assess a given executive function (van der Ven, 2011; van der Ven et al., 2012). This limitation should be addressed in future research.

Conclusions

Based on a Bayesian factor-analytic approach, our study extends previous findings by identifying various combinations of executive predictors of mathematical achievement in Grades 2, 6, and 10. Another contribution of this study concerns the quantification of the relative importance of executive variables. In this context, evidence is provided in favor of the differential contributions of various measures of inhibition, cognitive flexibility, and verbal WM to mathematical achievement across the three school grades. Third, logistic regression showed that the executive models derived from the Bayesian analyses had a similar ability to classify students with mathematical difficulties and their peers with typical mathematical achievement to the broader cognitive models that included fluid intelligence and processing speed. Finally, the study's main contribution was in identifying the different roles of EFs, fluid intelligence, and processing speed in predicting mathematical difficulties in each school grade.

Although further studies should be conducted to confirm these findings, the current results suggest considerations for the development of educational prevention and intervention programs within the curriculum framework (Clements & Sarama, 2019). First, it is important to compare the contributions of interventions focused on cognitive skills with those of interventions focused on specific domain skills and to determine their combined effect on academic achievement. In addition, it would be interesting to examine the preventive effects of promoting cognitive skills (such as verbal fluency, verbal working memory, and fluid intelligence), which were shown to be protective factors against mathematical difficulty in this study. Similarly, at a more specific level, it would be interesting to compare the effects of interventions focused on executive skills with those of interventions focused on strategic skills centered on the deficits shown in students with mathematical difficulties (e.g., Jitendra et al., 2009; Swanson, 2016).

Data availability

Data will be available at the Open Science Framework (OSF) and can be accessed at https://osf.io/vm4xe/?view_only=c11bed8575ce491d8ddcf15eb85bb77a.

Acknowledgment

We thank Mercedes Rucián and Laura Herrero for their help in data collections and the students and teachers who have collaborated with this research. We also thank Pedro R. Montoro for the task's desing. This work was financially supported by Project MINECO (EDU2011-22699).

Author contributions

Valentín Iglesias-Sarmiento: conceptualization, methodology, data analysis, writing–original draft, writing–review & editing; Nuria Carriedo: conceptualization, funding acquisition, project administration, data curation, writing–original draft, writing–review & editing; Odir A. Rodríguez-Villagra: methodology, data analysis; Leire Pérez: writing–review & editing.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jecp.2023.105715.

References

Agostini, F., Zoccolotti, P., & Casagrande, M. (2022). Domain-general cognitive skills in children with mathematical difficulties and dyscalculia: A systematic review of the literature. *Brain Sciences, 12*. https://doi.org/10.3390/brainsci12020239. Article 239.

Andersson, U. (2008). Mathematical competencies in children with different types of learning difficulties. *Journal of Educational Psychology, 100*(1), 48–66. https://doi.org/10.1037/0022-0663.100.1.48.

Andersson, U. (2010). Skill development in different components of arithmetic and basic cognitive functions: Findings from a 3-year longitudinal study of children with different types of learning difficulties. *Journal of Educational Psychology, 102*(1), 115–134. https://doi.org/10.1037/a0016838.

Archambeau, K., & Gevers, W. (2018). (How) are executive functions actually related to arithmetic abilities? In A. Henik & W. Fias (Eds.), *Heterogeneity of function in numerical cognition* (pp. 337–357). Academic Press. https://doi.org/10.1016/B978-0-12-811529-9.00016-9.

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology, 39*(1), 15–22. https://doi.org/10.1080/00221309.1948.9918159.

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.

Boschloo, A., Krabbendam, L., Aben, A., De Groot, R., & Jolles, J. (2014). Sorting Test, Tower Test, and BRIEF-SR do not predict school performance of healthy adolescents in preuniversity education. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00287. Article 287.

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology, 65*(1), 1–24. https://doi.org/10.1006/jecp.1996.2358.

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives, 8*(1), 36–41. https://doi.org/10.1111/cdep.12059.

Cai, D., Li, Q. W., & Deng, C. P. (2013). Cognitive processing characteristics of 6th to 8th grade Chinese students with mathematics learning disability: Relationships among working memory, PASS processes, and processing speed. *Learning and Individual Differences, 27*, 120–127. https://doi.org/10.1016/j.lindif.2013.07.008.

Camos, V., & Barrouillet, P. (2011). Developmental change in working memory strategies: From passive maintenance to active refreshing. *Developmental Psychology, 47*(3), 898–904. https://doi.org/10.1037/a0023193.

Carriedo, N., Corral, A., Montoro, P. R. P. R., Herrero, L., & Rucián, M. (2016). Development of the updating executive function: From 7-year-olds to young adults. *Developmental Psychology, 52*(4), 666–678. https://doi.org/10.1037/dev0000091.

Carriedo, N., & Rucián, M. (2009). Adaptación para niños de la prueba de amplitud lectora de Daneman y Carpenter (PAL-N) [Adaptation of Daneman and Carpenter's reading span test (PAL-N) for children]. *Infancia y Aprendizaje/Journal for the Study of Education and Development, 32*(3), 449–465. https://doi.org/10.1174/021037009788964079.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*(3), 386–404. https://doi.org/10.1016/0022-0965(82)90054-6.

Censabella, S., & Noël, M. P. (2007). The inhibition capacities of children with mathematical disabilities. *Child Neuropsychology, 14*, 1–20. https://doi.org/10.1080/09297040601052318.

Chan, B. M. Y., & Ho, C. S. H. (2010). The cognitive profile of Chinese children with mathematics difficulties. *Journal of Experimental Child Psychology, 107*, 260–279. https://doi.org/10.1016/j.jecp.2010.04.016.

Cirino, P. T., Fuchs, L. S., Elias, J. T., Powell, S. R., & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities, 48*(2), 156–175. https://doi.org/10.1177/0022219413494239.

Clements, D. H., & Sarama, J. (2019). Executive function and early mathematical learning difficulties. In A. Fritz, V. G. Haase, & P. Räsänen (Eds.), *International handbook of mathematical learning difficulties* (pp. 755–771). Springer. https://doi.org/10.1007/978-3-319-97148-3_43.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772.

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition, 162*, 923–931. https://doi.org/10.1016/j.cognition.2017.01.014.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6.

De Beni, R., & Palladino, P. (2004). The decline in working memory updating through ageing: Intrusion error analyses. *Memory, 12*(1), 75–89. https://doi.org/10.1080/09658210244000568.

Dempster, F. N. (1993). Resistance to interference: Developmental changes in a basic processing mechanism. In M. Howe & R. Pasnak (Eds.), *Emerging themes in cognitive development* (pp. 3–27). Springer. https://doi.org/10.1007/978-1-4613-9220-0_1.

Deng, M., Cai, D., Zhou, X., & Leung, A. W. S. (2022). Executive function and planning features of students with different types of learning difficulties in Chinese junior middle school. *Learning Disability Quarterly, 45*(2), 134–143. https://doi.org/10.1177/0731948720929006.

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

Ecker, U. K. H., Oberauer, K., & Lewandowsky, S. (2014). Working memory updating involves item-specific removal. *Journal of Memory and Language, 74*, 1–15. https://doi.org/10.1016/j.jml.2014.03.006.

Epstein, H. T. (2001). An outline of the role of brain in human cognitive development. *Brain and Cognition, 45*(1), 44–51. https://doi.org/10.1006/brcg.2000.1253.

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

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2009). Individual differences are almost entirely genetic in origin. *Journal of Experimental Psychology, 137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201.

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*, 29–44. https://doi.org/10.1016/J.EDUREV.2013.05.003.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory and fluid intelligence in children. *Biological Psychology, 54*(1–3), 1–34. https://doi.org/10.1016/S0301-0511(00)00051-X.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*(1), 30–47. https://doi.org/10.1037/0022-0663.100.1.30.

Geary, D. C., Hoard, M. K., Nugent, L., Ünal, Z. E., & Scofield, J. E. (2020). Comorbid learning difficulties in reading and mathematics: The role of intelligence and in-class attentive behavior. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.572099 572099.

Gerst, E. H., Cirino, P. T., Fletcher, J. M., & Yoshida, H. (2015). Cognitive and behavioral rating measures of executive function as predictors of academic outcomes in children. *Child Neuropsychology, 23*(4), 381–407. https://doi.org/10.1080/09297049.2015.1120860.

Gilmore, C., & Cragg, L. (2018). The role of executive function skills in the development of children's mathematical competencies. In A. Henik & W. Fias (Eds.), *Heterogeneity of function in numerical cognition* (pp. 263–286). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-811529-9.00014-5.

Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language, 92*, 183–201. https://doi.org/10.1016/j.jml.2016.06.004.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). Academic Press. https://doi.org/10.1016/b978-012208930-5/50007-6.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). MIT Press. https://doi.org/10.7551/mitpress/1480.003.0032.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*, 1166–1186. https://doi.org/10.3758/s13428-017-0935-1.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (2nd ed.). John Wiley.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010.

Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.

Jitendra, A. K., Star, J. R., Starosta, K., Leh, J. M., Sood, S., Caskie, G., Hughes, C. L., & Mack, T. R. (2009). Improving seventh grade students' learning of ratio and proportion: The role of schema-based instruction. *Contemporary Educational Psychology, 34*, 250–264. https://doi.org/10.1016/j.cedpsych.2009.06.001.

Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General, 145*, 1017–1048. https://doi.org/10.1037/xge0000184.

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

Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1339–1348. https://doi.org/10.1037/a0013078.

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *Journal of Clinical and Experimental Neuropsychology, 32*(5), 455–462. https://doi.org/10.1080/13803390903164363.

Lechuga, M. T., Moreno, V., Pelegrina, S., Gómez-Ariza, C. J., & Bajo, M. T. (2006). Age differences in memory control: Evidence from updating and retrieval-practice tasks. *Acta Psychologica, 123*(3), 279–298. https://doi.org/10.1016/j.actpsy.2006.01.00.

Lee, K., & Bull, R. (2016). Developmental changes in working memory, updating, and math achievement. *Journal of Educational Psychology, 108*(6), 869–882. https://doi.org/10.1037/edu0000090.

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology, 101*(2), 373–387. https://doi.org/10.1037/a0013843.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action. *Psychological Review, 91*, 295–327 https://psycnet.apa.org/record/1994-97487-005.

Lubin, A., Regrin, E., Boulc'h, L., Pacton, S., & Lanoë, C. (2016). Executive functions differentially contribute to fourth graders' mathematics, reading, and spelling skills. *Journal of Cognitive Education and Psychology, 15*(3), 444–463. https://doi.org/10.1891/1945-8959.15.3.444.

McAuley, T., & White, D. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology, 108*(3), 453–468. https://doi.org/10.1016/j.jecp.2010.08.009.

McDonald, P. A., & Berg, D. H. (2017). Identifying the nature of impairments in executive functioning and working memory of children with severe difficulties in arithmetic. *Child Neuropsychology, 24*(8), 1047–1062. https://doi.org/10.1080/09297049.2017.1377694.

McLean, J., & Hitch, G. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240–260. https://doi.org/10.1006/jecp.1999.2516.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734.

Mondloch, C. J., Geldart, S., Maurer, D., & de Schonen, S. (2003). Developmental changes in the processing of hierarchical shapes continue into adolescence. *Journal of Experimental Child Psychology, 84*(1), 20–40. https://doi.org/10.1016/S0022-0965(02)00161-3.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7.

Montoro, P. R., Luna, D., & Humphreys, G. W. (2011). Density, connectedness and attentional capture in hierarchical patterns: Evidence from simultanagnosia. *Cortex, 47*(6), 706–714. https://doi.org/10.1016/j.cortex.2010.05.007.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*(2), 111–121. https://doi.org/10.1111/j.2044-8295.1990.tb02349.x.

Mueller, S. T. (2011). *PEBL's Berg Card Sorting Test-64 (PBCST-64)* [computer software]. Retrieved from http://pebl.sf.net/battery.html.

Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B., & Early, M. C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities, 40*, 458–478. https://doi.org/10.1177/00222194070400050901.

Naglieri, J. A., & Das, S. (1997). *Das-Naglieri Cognitive Assessment System.* Riverside.

Ng, J., Lee, K., & Khng, K. H. (2017). Irrelevant information in math problems need not be inhibited: Students might just need to spot them. *Learning and Individual Differences, 60*, 46–55. https://doi.org/10.1016/j.lindif.2017.09.008.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220.

Nunes de Santana, A., Roazzi, A., & Nobre, A. P. M. C. (2022). The relationship between cognitive flexibility and mathematical performance in children: A meta-analysis. *Trends in Neuroscience and Education, 28*. https://doi.org/10.1016/j.tine.2022.100179 100179.

Paap, K. R., & Sawi, O. (2016). The role of test–retest reliability in measuring individual and group differences in executive functioning. *Journal of Neuroscience Methods, 274*, 81–93. https://doi.org/10.1016/j.jneumeth.2016.10.002.

Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *Journal of Clinical and Experimental Neuropsychology, 33*(4), 456–470. https://doi.org/10.1080/13803395.2010.533157.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition, 29*(2), 344–354. https://doi.org/10.3758/BF0319492.

Passolunghi, M. C., Cornoldi, C., & De Liberto, S. (1999). Working memory and intrusions of irrelevant information in a group of specific poor problem solvers. *Memory & Cognition, 27*, 779–790. https://doi.org/10.3758/bf03198531.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80*, 44–57. https://doi.org/10.1006/jecp.2000.2626.

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology, 88*, 348–367. https://doi.org/10.1016/j.jecp.2004.04.002.

Peng, P., Wang, C., & Namkung, J. (2018). Understanding the cognition related to mathematics difficulties: A meta-analysis on the cognitive deficit profiles and the bottleneck theory. *Review of Educational Research, 88*(3), 434–476. https://doi.org/10.3102/0034654317753350.

Pritchard, V. E., & Neumann, E. (2004). Negative priming effects in children engaged in nonspatial tasks: Evidence for early development of an intact inhibitory mechanism. *Developmental Psychology, 40*(2), 191–203. https://doi.org/10.1037/0012-1649.40.2.191.

Raven, J. C., Court, J. H., & Raven, J. (1996). *Manual for Raven's Progressive Matrices and Vocabulary Scales, Section 3: The Standard Progressive Matrices*. Oxford Psychologists Press.

Redick, T. S., Shipstead, Z., Meier, M. E., Montroy, J. J., Hicks, K. L., Unsworth, N., Kane, M. J., Hambrick, K. L., & Engle, R. W. (2016). Cognitive predictors of a common multitasking ability: Contributions from working memory, attention control, and fluid intelligence. *Journal of Experimental Psychology: General, 145*, 1473–1492. https://doi.org/10.1037/xge0000219.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 501–526. https://doi.org/10.1037/xlm0000450.

Rouder, J. N., & Morey, R. D. (2015). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research, 6*, 877–903. https://doi.org/10.1080/00273171.2012.734737.

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*(2), 573–594. https://doi.org/10.1207/s15326942dn2802_2.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*(4), 745–759. https://doi.org/10.1080/17470210500162854.

Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 1089–1096. https://doi.org/10.1037/a0015730.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Psychology Software Tools.

Simonds, J., Kieras, J. E., Rueda, M. R., & Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7–10-year-old children. *Cognitive Development, 22*(4), 474–488. https://doi.org/10.1016/j.cogdev.2007.08.009.

Skippen, P., Matzke, D., Heathcote, A., Fulham, W. R., Michie, P., & Karayanidis, F. (2019). Reliability of triggering inhibitory process is a better predictor of impulsivity than SSRT. *Acta Psychologica, 192*, 104–117. https://doi.org/10.1016/j.actpsy.2018.10.016.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651.

Swanson, H. L. (2016). Word problem solving, working memory and serious math difficulties: Do cognitive strategies really make a difference? *Journal of Applied Research in Memory and Cognition, 5*, 368–383. https://doi.org/10.1016/j.jarmac.2016.04.012.

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex, 49*, 2674–2688. https://doi.org/10.1016/j.cortex.2013.06.007.

The Jamovi Project (2021). *Jamovi* (Version 2.3.2) [computer software].

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology, 37*(4), 571–590. https://doi.org/10.1080/14640748508400920.

Uka, F., Gunzenhauser, C., Larsen, R. A., & von Suchodoletz, A. (2019). Exploring a bidirectional model of executive functions and fluid intelligence across early development. *Intelligence, 75*, 111–121. https://doi.org/10.1016/j.intell.2019.05.002.

Unsworth, N., & McMillan, B. D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Psychologica, 150*, 14–25. https://doi.org/10.1016/j.actpsy.2014.04.001.

Unsworth, N., Miller, A. L., & Robison, M. K. (2020). Are individual differences in attention control related to working memory capacity? A latent variable mega-analysis. *Journal of Experimental Psychology: General, 150*(7), 1332–1357. https://doi.org/10.1037/xge0001000.

van den Wildenberg, W. P. M., & van der Molen, M. W. (2004). Additive factors analysis of inhibitory processing in the stop-signal paradigm. *Brain and Cognition, 56*(2), 253–266. https://doi.org/10.1016/j.bandc.2004.06.006.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology, 87*(3), 239–266. https://doi.org/10.1016/j.jecp.2003.12.002.

van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology, 82*(1), 100–119. https://doi.org/10.1111/j.2044-8279.2011.02035.x.

Verbruggen, F., Aron, A. R., Band, G. P. H., Beste, C., Bissett, P. G., Brockett, A. T., Brown, J. W., Chamberlain, S. R., Chambers, C. D., Colonius, H., Colzato, L. S., Corneil, B. D., Coxon, J. P., Dupuis, A., Eagle, D. M., Garavan, H., Greenhouse, I., Heathcote, A., Huster, R. J., ... Boehler, C. N. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife, 8*. https://doi.org/10.7554/eLife.46323. Article e46323.

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods, 40*(2), 479–483. https://doi.org/10.3758/BRM.40.2.47.

Vukovic, R. K., & Siegel, L. S. (2010). Academic and cognitive characteristics of persistent mathematics difficulty from first through fourth grade. *Learning Disabilities in Research & Practice, 25*, 25–38. https://doi.org/10.1111/j.1540-5826.2009.00298.x.

Waszak, F., Li, S. C., & Hommel, B. (2010). The development of attentional networks: Cross-sectional findings from a life span sample. *Developmental Psychology, 46*(2), 337–349. https://doi.org/10.1037/a0018541.

Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K., & Pennington, B. F. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities, 46*(6), 500–516. https://doi.org/10.1177/0022219413477476.

Yeniad, N., Malda, M., Mesman, J., Marinus, H., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences, 23*, 1–9. https://doi.org/10.1016/j.lindif.2012.10.004.

Yuste, C., & Yuste, D. (2011). BADyG/E1: Batería de Aptitudes Mentales Diferenciales y Generales. *Nivel E1 [Battery of Differential and General Mental Aptitudes, Level E1]*. CEPE.
Yuste, C., Yuste, D., Martínez, R., & Galve, J. L. (2011). BADyG/E3: Battery of Differential and General Mental Aptitudes. *Nivel E3 [Battery of Differential and General Mental Activities, Level E3]*. CEPE.
Yuste, C., Yuste, D., Martínez, R., & Galve, J. L. (2012). *BADyG/M: Batería de Aptitudes Mentales Diferenciales y generales, Nivel M [Battery of Differential and General Mental Aptitudes, Level M]*. CEPE.
Zelazo, P. D., & Carlson, S. M. (2020). The neurodevelopment of executive function skills: Implications for academic achievement gaps. *Psychology & Neuroscience, 13*(3), 273–298. https://doi.org/10.1037/pne0000208.
Glass, B. D., Maddox, W. T., & Love, B. C. (2013). Real-time strategy game training: Emergence of a cognitive flexibility trait. *PLoS One, 8*(8). e70350. https://doi.org/10.1371/journal.pone.0070350.
Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental neuropsychology, 19*(3), 273–293. https://doi.org/10.1207/S15326942DN1903_3.
Davidson, M., Amso, D., Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*, 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006.
Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31.
Munro, S., Chau, C., Gazarian, K., & Diamond, A. (2006). Dramatically larger flanker effects (6-fold elevation). Cognitive Neuroscience Society Annual Meeting. In C. M. MacLeod (Ed.), *The Stroop Task in Cognitive Research* (pp. 17–40). In Cognitive methods and their application to clinical research. https://doi.org/10.1037/10870-002.
Somsen, R. J. M., Van Der Molen, M. W., Jennings, J. R., & Van Beek, B. (2000). Wisconsin Card Sorting in adolescents: Analysis of performance, response times and heart rate. *Acta Psychologica, 104*(2), 227–257. https://doi.org/10.1016/S0001-6918(00)00030-5.
Geary, D. C. (2013). Early Foundations for Mathematics Learning and Their Relations to Learning Disabilities. *Current Directions in Psychological Science, 22*(1), 23–27. https://doi.org/10.1177/0963721412469398.
Van der Ven, S.H.G. (2011). The structure of executive functions and relations with early math learning. (Doctoral dissertation, Utrecht University). doi.org/10.1111/j.2044-8279.2011.02035.x.