



Neurocognitive mechanisms explaining the role of math attitudes in predicting children's improvement in multiplication skill

Macarena Suárez-Pellicioni¹ · Ö. Ece Demir-Lira² · James R. Booth³

Accepted: 19 April 2021 / Published online: 6 May 2021
© The Psychonomic Society, Inc. 2021

Abstract

Enhancing student's math achievement is a significant educational challenge. Numerous studies have shown that math attitudes can predict improvement in math performance, but no study has yet revealed the underlying neurocognitive mechanisms explaining this effect. To answer this question, 50 children underwent functional magnetic resonance imaging (fMRI) when they were 11 (time 1; T1) and 13 (time 2; T2) years old. Children solved a rhyming judgment and a single-digit multiplication task inside the scanner at T1. The rhyming task was used to independently define a verbal region of interest in the left inferior frontal gyrus (IFG). We focused on this region because of previous evidence showing math attitudes-related effects in the left IFG for children with low math skill (Demir-Lira et al., 2019). Children completed standardized testing of math attitudes at T1 and of multiplication skill both at T1 and T2. We performed a cluster-wise regression analysis to investigate the interaction between math attitudes and improvement in multiplication skill over time while controlling for the main effects of these variables, intelligence, and accuracy on the task. This analysis revealed a significant interaction in the left IFG, which was due to improvers with positive math attitudes showing enhanced activation. Our result suggests that IFG activation, possibly reflecting effort invested in retrieving multiplication facts, is one of the possible neurocognitive mechanism by which children with positive math attitudes improve in multiplication skill. Our finding suggests that teachers and parents can help children do better in math by promoting positive math attitudes.

Keywords Attitudes · Math · Multiplication · Longitudinal · Children · functional magnetic resonance imaging

Introduction

Early education in the fields of science, technology, engineering, and mathematics (STEM) is crucial in preparing future workers for the demands for qualified employees in our society. Besides the importance of cognitive factors in predicting mathematics achievement (Geary, 2011), affective factors also play a role (Lipnevich et al., 2016). The affective domain in

mathematics education comprises attitudes, beliefs, and emotions (McLeod, 1992). Attitudes are defined as “learned predispositions or tendencies to respond positively or negatively to some object, situation, concept, or another person” (Aiken, 1970a). Attitudes towards math have been described as “an aggregated measure of liking or disliking of math, a tendency to engage in or avoid mathematical activities, a belief that one is good or bad at mathematics, and a belief that mathematics is useful or useless” (Neale, 1969, p. 632).

Correlational studies have shown widely varying associations between math attitudes and math achievement, from a meta-analysis reporting a significant but weak effect size, without accounting for differences in math skill (Xin Ma & Kishor, 1997), to other studies reporting that math attitudes accounted for approximately 26% of the variance in arithmetic performance in a group of first graders (Dowker et al., 2019) or that math attitudes explained up to 25% of the variance in math grades even after controlling for math skill (Lipnevich et al., 2016). Studies have shown that the association between math attitudes and math achievement remains significant even

✉ Macarena Suárez-Pellicioni
mspelllicioni@ua.edu

¹ Department of Educational Studies in Psychology, Research Methodology, and Counseling, University of Alabama, 270 Kilgore Ln, Tuscaloosa, AL, USA

² Department of Psychological and Brain Sciences, DeLTA Center, Iowa Neuroscience Institute, The University of Iowa, Iowa City, IA, USA

³ Department of Psychology and Human Development, Vanderbilt University, Nashville, TN, USA

after accounting for age, IQ, working memory, and math anxiety (Chen et al., 2018). Correlational studies are limited because they only show an association between variables, and do not allow an inference regarding the directionality of the effects. The question that teachers, parents, and researchers find of educational relevance is whether having math attitudes early on can explain children's improvement in math over time. This question is particularly important, because, unlike some variables that impact math achievement, such as home backgrounds or school characteristics, attitudes are relatively malleable and amenable to change (Singh et al., 2002). Therefore, modifying attitudes could be a valuable tool in helping children do better in math. Mazzocco et al. (2012) showed that early math attitudes can explain later math performance even in young children. They found that second-graders' beliefs about math likeability, difficulty, and usefulness, as expressed by children's spontaneous conversations, predicted children's math calculations scores in third grade. However, this study did not control for initial levels of children's math skill.

Controlling for initial levels of the dependent measure in a longitudinal study is crucial for addressing the role that variable "x" has in predicting *improvement* in variable "y" over time. As claimed by Merkle and colleagues, controlling for initial levels of the dependent measure is important, because it is very likely that the differences found in the dependent measure at time 2 (T2; i.e., math achievement at T2) are due to initial differences in that variable at time 1 (T1; i.e., math achievement at T1) instead of to the effect of the predictor of interest (i.e., math attitudes at T1) (Merkley et al., 2017). Many longitudinal studies have shown that children's early math skill strongly predict later math achievement (Aubrey et al., 2006; Reynolds, 1991; Yates, 2000). Therefore, it is crucial to determine whether math attitudes explain children's math improvement regardless of their initial levels of math performance.

Several longitudinal studies have found that math attitudes predict improvement in math skill over and above initial levels of math performance. Ai (2002) found that seventh graders' math attitudes predicted improvement in math performance from seventh to tenth grade. Hemmings et al. (2011) found that although 12-year-old's math ability at T1 was a significant predictor of math exam scores 3 years later, adding that math attitudes at T1 as a predictor in the model resulted in a larger predictive capacity. Similar results were obtained by studies looking more broadly at the role of attitudes toward school in predicting math improvement. Positive school attitudes have been found to predict math improvement across all four waves of a longitudinal study, from kindergarten to third grade (Bodovski & Farkas, 2007) and from third to sixth grade (Borman & Overman, 2004). Studies also have reported that liking and showing interest in math, which is an important component of math attitudes, predicted how much children

improved in math. Aunola et al. (2006) measured the liking of math and math performance at the beginning (T1) and end (T2) of first grade and once again in second grade (T3). They found a direct effect of liking of math at T2 on math performance at T3, even after accounting for initial levels of math performance. When comparing changes in interest and changes in math performance they found that the more children liked math over time, the more they improved in math. In a study with even younger children, Fisher et al. (2012) found that 4-year-old children's interest in math in the fall (T1), defined as the eagerness and sophistication when playing with math toys, predicted later math achievement in the spring (T2), even after controlling for earlier math skill. The same eagerness to learn predicted children's longitudinal improvement in math achievement from kindergarten to third grade in the DiPerna et al. (2007) study. Overall, this evidence corroborates the well-known idea that we learn better what we like and find interesting and that promoting interest among students is an excellent way to get them engaged, motivated, and to help them excel in their learning (Harackiewicz et al., 2016).

Besides math liking and interest, other crucial components of math attitudes are perceptions of one's math skills and perceptions of the usefulness of math. According to expectancy-value theory (Wigfield & Eccles, 2000), individuals' choice, persistence, and performance in a given domain can be explained by their beliefs about how well they will do on the activity (i.e., self-perceptions of skill) and the extent to which they value the domain (i.e., perceived usefulness). Numerous studies have supported this theory by showing that children who perceive themselves as good at math and that math is important to their future are more likely to pursue advanced math classes than students who perceive themselves as bad at math or who think that math is useless (Crombie et al., 2005; Updegraff et al., 1996). Self-perceptions of one's math skills have been shown not only to be associated with math performance (Dowker et al., 2019) but

to predict improvement longitudinally in math skill. Several studies have found that children's self-perceptions of their math abilities in fifth grade (Petersen & Hyde, 2017), at the beginning of eighth grade (Trautwein et al., 2009), and in first through fourth grade (Wigfield & Eccles, 2000), predict math achievement 5 years later at the end of eighth grade and in high school, respectively, even after accounting for initial levels of performance. Although perceptions of the usefulness of math have been shown not to directly predict improvement in math over and above earlier math performance (Reynolds & Walberg, 1992), it was reported to be the strongest predictor of children's intentions and actual decision to keep taking math courses (Wigfield & Eccles, 2000), which would eventually impact math performance. All of this evidence shows that regardless of whether math attitudes are studied as a whole (Ai, 2002; Hemmings et al., 2011), or specifically

focusing on its interest (Aunola et al., 2006; DiPerna et al., 2007; Fisher et al., 2012) and perceptions (Petersen & Hyde, 2017; Trautwein et al., 2009; Wigfield & Eccles, 2000) components, they predict *changes* in math performance over time at different stages of formal math development.

To the best of our knowledge, only two studies have investigated the underlying neurocognitive mechanisms of the association (Chen et al., 2018) and the interaction (Demir-Lira et al., 2019) between math attitudes and math performance. First, using functional magnetic resonance imaging (fMRI), Chen et al. (2018) investigated 7-to-11-year-old children while they solved single-digit additions. Using structural equation modeling, they found that the effects between positive math attitudes and math achievement were mediated by increased use of retrieval strategy and greater hippocampal activation, a region involved in learning and memory formation. However, the authors acknowledged that because of the cross-sectional nature of their study, they were unable to determine the directionality of the effects. Second, our fMRI study in 8- to 15-year-old children solving multiplication problems examined the neurocognitive mechanisms explaining the interaction between math skill and math attitudes but did not consider change over time (Demir-Lira et al., 2019). Our regression analysis showed a significant interaction between math attitudes and math skill in the left inferior frontal gyrus (IFG), which was due to children with low skill but positive math attitudes showing greater activation in this region compared with their low skilled peers with negative math attitudes. Previous studies had suggested that the IFG is involved in the access to arithmetic facts stored in long-term memory (Prado et al., 2011), as it shows less activation as fact retrieval becomes more automatic over development (Prado et al., 2014; Rivera et al., 2005) and shows greater activation when the math task is more effortful (Delazer et al., 2003; Soltanlou et al., 2017; Suárez-Pellicioni et al., 2018). Based on these studies, we interpreted our results as showing that children with low skill and positive math attitudes were exerting more effort in retrieving multiplication facts.

Given that previous neuroimaging studies examining the relation of math attitudes and achievement were cross-sectional, they could not determine the neurocognitive mechanisms involved in predicting math gains. The current study examined whether enhanced IFG activation associated with positive math attitudes predicts greater improvement in math over time, even after accounting for initial levels of skill. We followed some of the participants in the Demir-Lira et al. (2019) study longitudinally, which allows inferences about the directionality of the effects between math attitudes and math improvement. Our analytical approach focused on a verbal region of interest (ROI) in left IFG suggested in this study to be related to the effort expended on solving math problems. If greater activation in left IFG for those with positive attitudes predicts gains in multiplication skill, then we can be more

confident that this proposed neurocognitive mechanism explains subsequent changes in math performance.

Methods

Participants

Sixty-five children from third to eighth grade were recruited from schools in the Chicago metropolitan area to participate in this study. This longitudinal dataset has been deposited in OpenNeuro (10.18112/openneuro.ds001486.v1.1.0) and a detailed description of it is given in Suárez-Pellicioni, Lytle, Younger, & Booth, (2019). Time point 1 of this dataset is the basis of other publications by our research group, including Berteletti et al. (2014), Berteletti and Booth (2015), Demir, Prado, & Booth (2014), Demir et al. (2015) Prado et al. (2014). The longitudinal data of this dataset is the basis of other publications, including Suárez-Pellicioni and Booth (2018), Suárez-Pellicioni et al. (2018), Suárez-Pellicioni, Fuchs, and Booth (2019), Suárez-Pellicioni et al. (2020), and Demir-lira et al. (2016). None of these studies have looked at the role of attitudes at T1 in predicting longitudinal improvement in multiplication skill over time, which constitutes the objective of this study.

All participants were native English speakers, right-handed, were free of past and present psychiatric disorders, including attention deficit hyperactivity disorder (ADHD), neurological disease, or epilepsy. According to parental report, no participant had hearing impairments, uncorrected visual impairment, was born prematurely (less than 36 weeks), was taking medication affecting the central nervous system, or had any contraindication for being scanned, such as having braces. Participants had no history of intellectual or reading deficits, all of them scoring above 80 standard scores (hereinafter, SS) on the Full IQ scale of the Wechsler Abbreviated Scale of Intelligence – WASI (Wechsler, 1999) and above 80 SS on the Word Attack subtest from the Woodcock-Johnson III Test of Achievement (WJ-III; Woodcock et al., 2001), respectively. Children and their parents or guardians provided written consent to participate in the study. Parents were compensated \$20 per hour for their time. All experimental procedures were approved by the Institutional Review Board at Northwestern University.

Fifteen of the initial 65 participants had to be excluded from the final sample for the following reasons: 1) being left-handed ($n = 1$); 2) excessive movement in the multiplication task solved inside the scanner ($n = 2$) where excessive movement was defined as more than 10% of the total volumes replaced or more than five consecutive volumes replaced in a given run; 3) scoring below 80 SS in the Word Attack subtest from the WJ-III (Woodcock et al., 2001; $n = 2$); 4) missing data from the Comprehensive Mathematical Abilities Test

(Hresko et al., 2003) at T2 ($n = 2$); 5) accuracy below 50% for the small ($n = 2$) or control ($n = 4$) conditions of the multiplication task solved inside the scanner; 6) problem with the program presenting the stimuli (i.e., Eprime) ($n = 2$). The final sample consisted of 50 participants¹ who were tested longitudinally, with sessions being approximately 2 years apart. The level of education of the children's parents was relatively high, being predominantly graduate (15 mothers and 14 fathers), bachelor's (13 mothers and 11 fathers), and some college (15 mothers and 10 fathers). A limited number of families completed high school (2 mothers and 7 fathers) or did not complete high school (1 mother and 2 fathers).² More detailed information about the sample is given in Table 1.

Standardized measures

The following standardized tests were used to assess different cognitive and attitudinal variables that constituted the covariates of interest, covariates of no interest, or variables that we used to rule out cognitive deficits before including participants in the final sample.

Math attitudes

Children's math attitudes were measured by the math attitudes subtest of the Test of Mathematical Abilities—Third Edition (TOMA-3; Brown et al., 2012). The math attitudes subtest includes 15 items related to children's liking or disliking of math (e.g., "I've always liked math"), their tendency to engage in or avoid math (e.g., "I'd rather do math than any other kind of homework"), their self-perception of math skill (e.g., "Math tests are usually easy for me"), and how useful they thought math is (e.g., "I use math a lot outside of school"). Children were presented with these 15 sentences and were asked to choose among four options ranging from "yes definitely" to "no definitely." Raw scores range from 15 to 60. Higher scores indicate more positive attitudes towards math. More detailed information about math attitudes raw scores is given in Table 1. Raw scores in this test at T1 were included in the cluster-wise regression analysis described in section 2.7.3, as a covariate of no interest (i.e., control of the main effect of math attitudes).

Multiplication skill

Multiplication skill was measured by the multiplication subtest of the Comprehensive Mathematical Abilities Test (CMAT; Hresko, Schlieve, Herron, Swain, & Sherbenou,

Table 1 Sample's information. Means (standard deviation in parenthesis) for age, time between sessions, and scores on standardized tests of math attitudes, multiplication skill, improvement, reading skill, intelligence (IQ), and accuracy (percentage) and response times (RTs; sec) for small and large multiplication problems solved inside the scanner, for the whole sample ($n = 50$)

	Whole sample ($n = 50$)
Age at T1 session (yr)	11.1 (1.5)
Age at T2 session (yr)	13.3 (1.6)
Time between sessions (yr)	2.2 (0.3)
Female/male ratio	26/24
Math attitudes at T1 ^a	42.8 (8.3)
Multiplication skill T1 ^a	11.8 (5.6)
Multiplication skill T2 ^a	14.0 (5.7)
Improvement T2-T1 ^a	0.0 (5.2)
Reading skill at T1 ^b	105.2 (10.8)
Full IQ at T1 ^b	112.3 (14.7)
% Accuracy: small	89.5 (12.8)
% Accuracy: large	68.0 (17.1)
RTs (sec): small	1.16 (.36)
RTs (sec): large	1.37 (.39)

^a Raw scores; ^b Standard scores. T1: Time 1; T2: Time 2. See section 2.2 for a description of the standardized tests.

2003), including 26 items. This is an untimed test including a wide range of multiplication problems, from very simple ones, presenting multiplication problems in non-symbolic format, up to more difficult ones including the multiplication of 4- x 3-digit numbers (e.g. 5,764 x 678) and the multiplication of fractions. Raw scores ranged from 0 to 26. This test was administered both at T1 and at T2. Scores in this test were used to calculate improvement over time in multiplication skill. This improvement measure was included as a covariate of no interest in the cluster-wise regression analysis described in section 2.7.3. (i.e., control of the main effect of improvement).

Intelligence

Full IQ, including both performance and verbal IQ, was measured by the Wechsler Abbreviated Scale of Intelligence (WASI; Weschler, 1999). Performance IQ was measured with the Block Design and the Matrix Reasoning subtests of the WASI. The Block Design subtest requires participants to use red-and-white blocks to re-create the model design shown by the examiner, within a specified time limit. In the Matrix Reasoning subtest, the examinee views an incomplete matrix or series and selects the response option that completes the matrix or series. Verbal IQ was measured with the Vocabulary and Similarities subtests of the WASI. In the Vocabulary subtest, the participant has to define words, while in the

¹ The following participants were included in this study: 2 5 6 7 8 9 10 11 12 16 20 22 23 24 27 34 35 36 40 44 45 46 48 49 50 53 54 55 56 57 60 61 65 66 67 69 70 71 73 74 75 76 86 89 90 93 95 96 103 106.

² Information was missing for the level of education of 4 mothers and 6 fathers.

Similarities subtest the participants are presented with two words that represent common objects or concepts and they have to describe how they are similar. Full IQ at T1 was used as a covariate of no interest in the cluster-wise regression analysis described in section 2.7.3.

Reading skill

Reading skill was measured by the Word Attack subtest from the Woodcock-Johnson III Test of Achievement (WJ-III; Woodcock et al., 2001), which requires reading phonologically regular nonsense words. This measure was used to make sure that participants did not have reading deficits, with all participants included in the final sample scoring above 80 SS in this test.

Scanner tasks

Rhyming judgment localizer task to identify verbal regions in frontal cortex

In the rhyming judgment task, two written monosyllabic English words were sequentially presented. Participants were asked to decide whether the words rhymed or not. To ensure that participants did not rely on orthography to solve the task, four conditions were created in which pairs of words had: 1) similar orthography and similar phonology (i.e., O+P+; e.g., *dime–lime*; 12 trials); 2) similar orthography but different phonology (i.e., O+P-; e.g., *pint–mint*; 10 trials); 3) different orthography but similar phonology (i.e., O-P+; e.g., *jazz–has*; 10 trials); 4) different orthography and different phonology (i.e., O-P-; e.g., *press–list*; 14 trials). The O+P+ and O-P- constituted the nonconflicting conditions, given that orthographic information was consistent with the right answer, whereas the O-P+ and O+P- conditions constituted the conflicting conditions because orthographic information was inconsistent with the correct answer. Including these conditions ensured that participants could not do the task based on visual information alone, but rather had to access the pronunciation. Figure 1a shows an example of the O-P+ condition of the rhyming judgment task and its timing. The control condition consisted of a blue square that was presented for the same duration as the experimental conditions. For the control condition, shown in Fig. 1b, children were asked to press a button when the square turned red. All participants received stimuli in the same order and in a single run, which lasted approximately 7 minutes.

Experimental task: Single-digit multiplication verification task

Participants were presented with two runs of a single-digit multiplication verification task. Each run comprised 12 small

problems and 12 large problems. Operands were smaller or equal to 5 for small problems and larger than 5 for large problems. The proposed solution could be correct or incorrect. Incorrect proposed solutions were the result of multiplying the first operand plus or minus 1 and the second operand (e.g., $6 \times 3 = 15$ or 21, which are the solutions for 5×3 and 7×3 , respectively). Problems involving 0 or 1 and ties (e.g., 3×3) were not included in the main experiment but were used in the practice session.

Each run consisted of 36 problems: 12 small problems with correct solutions, 12 large problems with correct solutions, 6 small problems with false solutions, and 6 large problems with false solutions.³ Figures 2C and D show an example of small and large multiplications, respectively, and their timing. A total of 72 problems were presented in two runs. Participants' task was to decide, employing button press, whether the proposed solution was correct or incorrect. Each run also included 12 control trials, which consisted of a blue square that was presented for the same time as the stimulus in the experimental conditions, and participants were asked to press the response button with their index finger when it turned red (Fig. 1b). The duration of each run was approximately 4 minutes. Twelve problems with a correct proposed solution and 12 problems with an incorrect proposed solution were included in the practice session. Different sets of stimuli were used in the practice and the scanning sessions.

Experimental protocol

Subjects participated in a practice session after informed consent was obtained and standardized tests were administered. During this session, they practiced all trials and learned to minimize head movement in a mock fMRI scanner. The actual scanning session took place within a week of the practice session. In the fMRI scanner, participants performed two runs of the multiplication task and one run of the rhyming judgment task. Stimuli were projected onto a screen that was viewed by the participants through a mirror attached to the head-coil. The order of the tasks was counterbalanced across participants. The timing and order of trial presentation within each run was optimized for estimation efficiency using optseq (<http://surfer.nmr.mgh.harvard.edu/optseq/>). Behavioral responses were recorded using an MR-compatible keypad

³ We included a greater proportion of problems with true than false solutions because we anticipated that false problems would be associated with error detection and conflict monitoring mechanisms. Despite the different proportion of problems with correct and incorrect proposed solutions, we did not observe any bias towards answering “yes” in the task, with 90% of the participants answering above chance level to problems with incorrect proposed solutions (i.e., which required a “no” answer). As indicated in section 2.7.2., given that problems with correct and incorrect proposed solutions showed no differences in activation in the region of interest, the two problems were analyzed together in the fMRI analysis to increase statistical power and equate number of items across time and improvement.

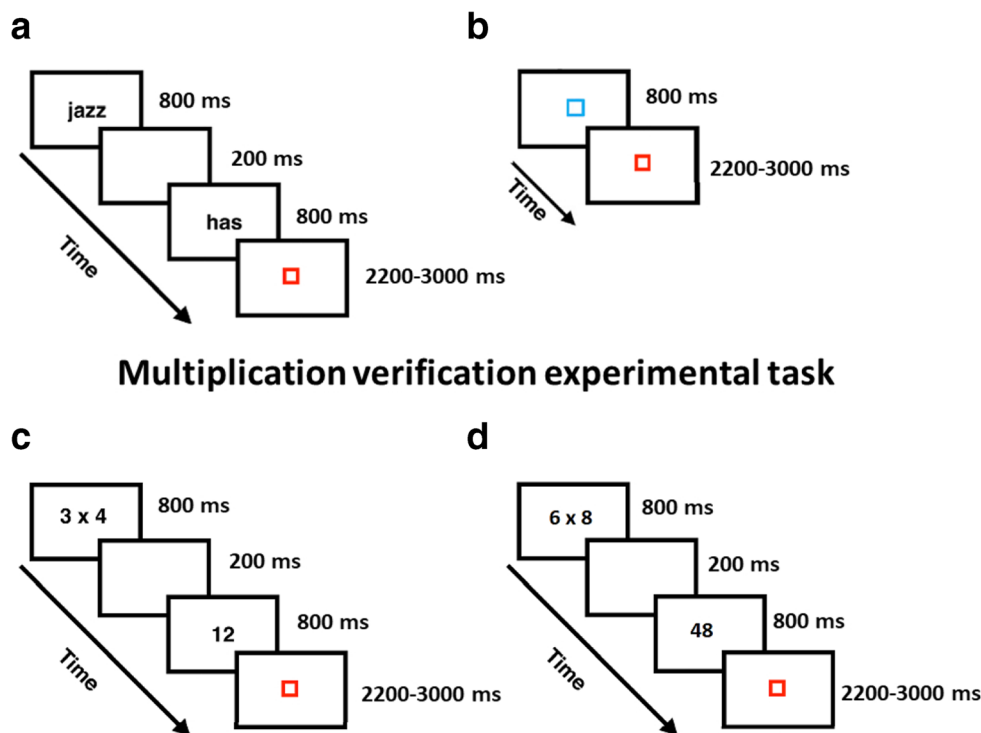


Fig. 1 Tasks solved inside the scanner. (a) The rhyming judgment task was used to localize verbal regions of the brain in inferior frontal cortex. In this task, participants had to respond to whether pairs of words rhymed or not. (b) The control condition common to all tasks solved inside the

scanner, in which participants had to press a button when the blue square turned red. Multiplication task: Single-digit multiplication verification task, including (c) small and (d) large multiplication problems

and participants responded with their right hand. In the multiplication task, participants responded with their index finger if the proposed solution was correct and with the middle finger if the proposed solution was incorrect. As for the rhyming judgment task, participants responded with their index finger if the words rhymed and with the middle finger if the words did not rhyme.

Stimulus timing

Stimulus timing was identical in all tasks. A trial started with the presentation of a first stimulus (i.e., multiplication problem or first word, depending on the task) for 800 ms, followed by a blank screen for 200 ms. A second stimulus (i.e., multiplication proposed solution or second word, depending on the task) was then presented for 800 ms, followed by a red fixation square for 200 ms. Variable periods of fixation, lasting 2,200, 2,600, or 3,000 ms (i.e. 400 ms jitter), were added after each trial, during which the red square was presented. Each run ended with 22 s of passive visual fixation to help with the deconvolution of the final trials.

fMRI acquisition

Images were collected using a Siemens 3T TIM Trio MRI scanner (Siemens Healthcare, Erlangen, Germany) at

Northwestern University's Center for Advanced MRI. The fMRI blood oxygenation level dependent (BOLD) signal was measured with a susceptibility weighted single-shot echo planar imaging (EPI) sequence. The following parameters were used: TE = 20 ms, flip angle = 80°, voxel size: 1.7 x 1.7 x 3 mm, matrix size = 128 x 120 x 37, field of view = 220 x 206.25 x 111 mm, slice thickness = 3 mm (0.48 mm gap), number of slices = 32, TR = 2,000 ms. Before functional image acquisition, a high resolution T1 weighted 3D structural image was acquired for each subject, with the following parameters: TR = 2,300 ms, TE = 3.36 ms, matrix size = 256 x 256, field of view = 240 mm, slice thickness = 1 mm, number of slices = 160.

fMRI data analysis

Preprocessing

Data analysis was performed using SPM12 (www.fil.ion.ucl.ac.uk/spm). The first six images of the run were discarded to allow for T1 equilibration effects. The remaining functional images were corrected for slice acquisition delays and realigned to their mean functional image across runs. The anatomical image was then segmented and warped to the default tissue probability map (TPM) template to get the transformation field. An anatomical brain mask was created by

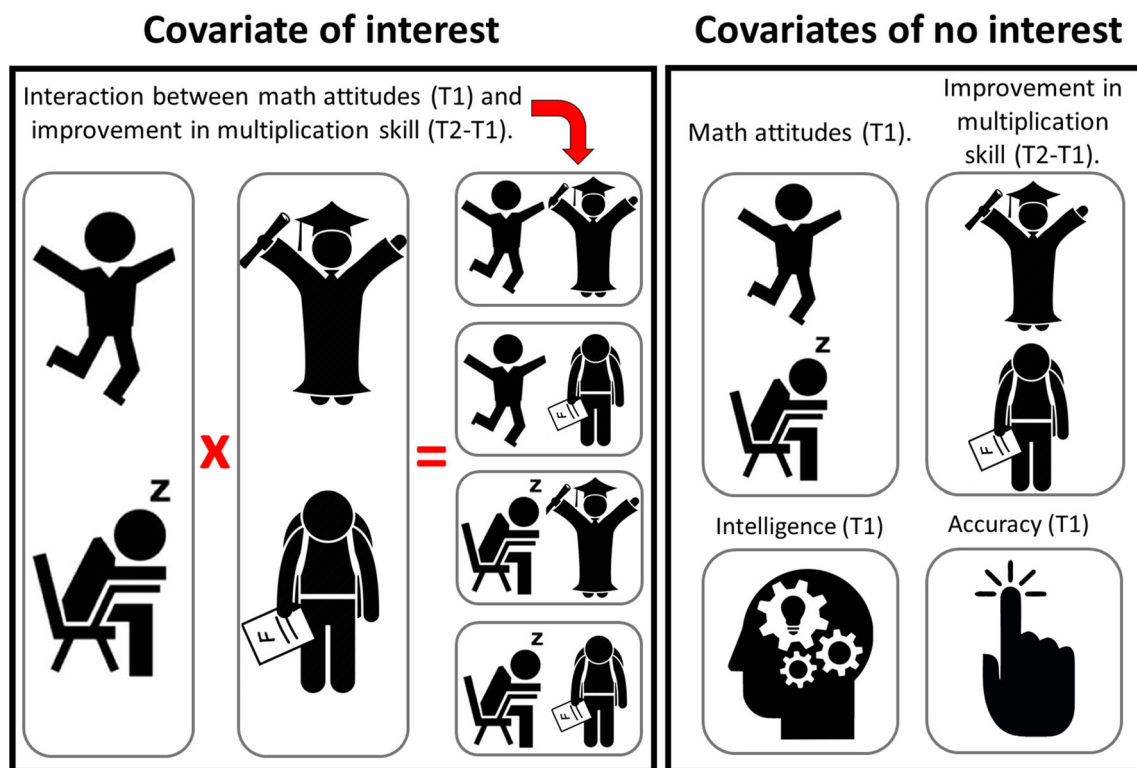


Fig. 2 Covariates in the cluster-wise regression analyses. Illustration of the covariates of interest and covariates of no interest (i.e., controls) included in the two cluster-wise regression analyses carried out to study the association between math attitudes and improvement in multiplication skill while children solved small and large multiplication problems inside the scanner. All variables were continuous. For the two regressions, the covariate of interest was the interaction between math attitudes and improvement in multiplication skill over time. The main effects of math

attitudes, the main effect of improvement in multiplication skill, and children's full IQ were included as covariates of no interest in the two regressions. When we studied this interaction with small multiplication task solving, we included accuracy in small multiplication problems solved inside the scanner as the fourth covariate of no interest. When the interaction was studied with large multiplication problems inside the scanner, accuracy in solving large problems was included as the fourth covariate of no interest.

combining the segmentation products (i.e., grey, white and cerebrospinal fluid), and then applied to its original anatomical image to produce a skull-stripped anatomical image. After that, the mean functional image and all functional images were co-registered to the skull-stripped anatomical image. Images were then normalized to the standard T1 Montreal Neurological Institute (MNI) template and spatially smoothed with a Gaussian filter equal to twice the voxel size ($4 \times 4 \times 8 \text{ mm}^3$ full width at half maximum). To reduce movement effects on brain signal, Art-Repair (<https://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html>; Mazaika et al., 2009) was used to identify outlier volumes, defined as those with volume-to-volume head movement exceeding 1.5 mm in any direction, head movement greater than 5 mm in any direction from the mean functional image across runs, or deviations of more than 4% from the mean global signal intensity. The outlier volumes were repaired by interpolation between the nearest nonoutlier volumes. Subjects included in our study had no more than 10% of the volumes repaired in each run and no more than 5 consecutive volumes repaired in each run. Six motion parameters estimated in the

realignment step were entered in the first level modeling as regressors and the repaired volumes were deweighted (Mazaika et al., 2009).

fMRI analysis: First-level

Event-related statistical analysis was performed according to the general linear model. Activation was modeled as epochs with onsets time-locked to the presentation of the first stimulus in each trial. To equate for power in the analysis and because no differences in activation were found between problems with a correct or incorrect proposed solution, all children's responses (i.e., correct and incorrect) were included in the model. All epochs were convolved with a canonical hemodynamic response function. The time series data were high-pass filtered (1/128 Hz), and serial correlations were corrected using an autoregressive AR model.

fMRI analysis: Second-level

We evaluated the interaction between math attitudes and improvement in multiplication skill over time while children solved multiplication problems inside the scanner. Given previous evidence reporting an interaction of math attitudes by math skill only for small multiplication problems (Demir-Lira et al., 2019), we studied this interaction separately for small and large multiplication problems, using the “small vs. control” and “large vs. control” contrasts, respectively. This resulted in two regression analyses, one for each problem size. First, we studied brain activation during small multiplication problems by carrying out a cluster-wise regression analysis, including the interaction between math attitudes and improvement in multiplication skill, as the covariate of interest. The main effects of attitudes and improvement in multiplication skill were included as covariates of no interest (i.e., controls). Given that improvement groups showed significant differences in full IQ and in accuracy for small multiplication problems solved inside the scanner (Table 4), these two measures were also included as covariates of no interest in the model. Second, the same cluster-wise regression analysis was carried out to study the interaction between math attitudes and improvement in multiplication skill while children solved large multiplication problems, but in this case, we included accuracy on large multiplication problems solved inside the scanner as the covariate of no interest. All these measures were continuous variables. Figure 2 shows an illustration of the covariates of interest and no interest included in the regression analyses.

Definition of the region of interest (ROI)

To localize verbal regions in the brain we identified, for each participant, the clusters that showed greater activation for all word pairs of the rhyming judgment task compared with the control condition, at T1. In a second-level analysis, these individual contrasts were submitted to a one-sample *t*-test across all participants. Given that the study by Demir-Lira et al. (2019) showed a cluster in left IFG being associated with the interaction between math attitudes and math skill in a multiplication task (Demir-Lira et al., 2019) and the extensive evidence that suggested that the left IFG is responsible for accessing verbal representations (Andin et al., 2015; Bookheimer, 2002; Booth et al., 2003, 2004; Fedorenko et al., 2012; Poldrack et al., 1999; Pollack & Ashby, 2017; Prado et al., 2011, 2014; Rickard et al., 2000), we constrained the brain activation elicited by the contrast “all word pairs vs. control” of the rhyming task within the anatomical left IFG.

Statistical significance for this anatomical left IFG was determined using Monte Carlo simulations in AFNI’s 3dClustSim program (December, 2015; see <http://afni.nimh.nih.gov/>, with SPM’s data smoothness parameters,

autocorrelation function [ACF] = 0.45, 4.14, 11.02). 3dClustSim carries out a user-specified number of Monte Carlo simulations of random noise activations at a particular voxel-wise alpha level within a masked brain volume. Following the suggestions made by Eklund et al. (2016) regarding the inflated statistical significance achieved using some packages (i.e., SPM, FSL, and AFNI), we used 3dClustSim’s most recent version (December 2015). We used 3dFWHMx to calculate the smoothness of the data for every participant, using a spatial ACF, and then averaged those smoothness values across all participants. This averaged smoothness value was then entered into 3dClustSim to calculate the cluster size needed for significance for a given anatomical mask. A cluster size of 46 was needed to reach significance for the anatomical left IFG region. Clusters exceeding this threshold, at a cluster-wise threshold of $p = 0.05$ and voxel-wise threshold of $p = 0.005$, were deemed significant.

The combination of the functional contrast “all words vs. control” within the anatomical left IFG resulted in a cluster of 3668 voxels, shown in Fig. 3. More information about this cluster is given in Table 2.

We then ran the 3dclustSim estimation for this combined ROI to determine the cluster size needed for significance within this region ($k = 3668$). A cluster size of 38 voxels was needed for significance, at a cluster-wise threshold of $p = 0.05$ and voxel-wise threshold of $p = 0.005$.

Whole-brain level analysis

To investigate nonpredicted effects outside the ROI in frontal cortex, we also performed a whole-brain analysis. We looked at clusters outside our left IFG ROI showing activation for the regression analyses described in section 2.7.3. and depicted in Fig. 2. Statistical significance for the whole-brain was also defined using 3dClustSim. A cluster of at least 202 voxels was needed to reach significance at the whole-brain level, at corrected level cluster-wise threshold of $p = 0.05$ and voxel-wise threshold of $p = 0.005$.

Results

Association between math attitudes and other variables

There were no sex-related differences in math attitudes at T1 ($t(48) = 0.35$, $p = 0.73$) or T2 ($t(48) = -0.66$, $p = 0.51$). Math attitudes at T1 showed a significant positive correlation with math attitudes at T2 ($r = 0.63$, $p < 0.001$) and with multiplication performance at T1 (i.e., standardized test; $r = 0.37$, $p = 0.008$), but not with any of the following variables: age at T1 ($r = 0.10$, $p = 0.51$), age at T2 ($r = 0.09$, $p = 0.54$),

Table 2 Localizer task contrast and anatomical constraint used to create the combined region of interest (ROI) implicated in verbal processing

Localizer contrast	Anatomical constraint	K	aal	~BA	MNI coordinate			Z value
					X	Y	Z	
All word pairs > control	Left IFG	3668	IFG Operularis, IFG Triangularis, IFG Orbitalis	44, 45, 46, 47	-46	10	24	Infinite
					-48	34	10	Infinite
					-46	28	16	6.92

Detailed information of the combined ROI including cluster size (K), corresponding region based on anatomical automatic labeling (aal), approximate Brodmann areas (~BA), MNI coordinates of the peaks, and Z values.

multiplication performance at T2 ($r = 0.23$, $p = 0.11$), improvement in multiplication skill ($r = 0.09$, $p = 0.55$), IQ ($r = 0.22$, $p = 0.13$), reading skill ($r = -0.04$, $p = 0.76$), or performance on the multiplication task solved inside the scanner (all p values > 0.11). We did not observe any tendency of math attitudes becoming more negative with age in our sample, the two variables were not significantly correlated ($r = 0.10$, $p = 0.51$).

Behavioral results: Performance on the rhyming judgment localizer task

We calculated accuracy (i.e., percentage of correctly solved trials) and response times (i.e., for correctly solved trials) for the rhyming judgment localizer task solved inside the scanner, for every participant. We then ran a repeated-measures ANOVAs including Conflict and Rhyming as two within-subjects factors. Conflict referred to whether orthography was consistent (i.e., nonconflicting) or inconsistent (i.e., conflicting) with the correct answer. Rhyming referred to whether word pairs rhymed or not. Post-hoc tests, using Bonferroni correction, were calculated when an effect was found significant.

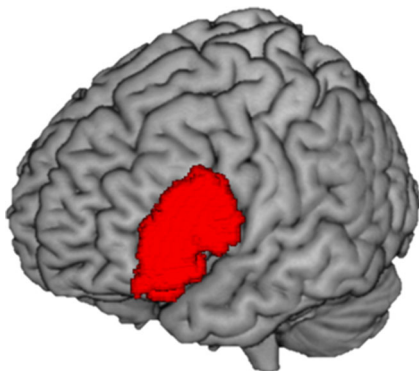


Fig. 3 Region of interest resulting from constraining brain activation elicited by comparing all word pairs of the rhyming judgment localizer task with the control condition within the anatomical left inferior frontal gyrus (IFG)

The Rhyming by Conflict interaction was significant for accuracy ($F(1,49) = 44.71$, $p < 0.001$, partial $\eta^2 = 0.47$), and it was due to significant differences between all the word pairs (all above $p = 0.03$) except between O-P+ and O-P- ($p = 1$). The easiest condition in terms of accuracy was the O+P+ (mean = 94.8, SEM = 1.2) whereas the hardest condition was the O+P- (mean = 57.0, SEM = 4.0) consisting of word pairs that did not rhyme and for which orthography did not help solve the task. Accuracy was intermediate for the conditions O-P+ (mean = 87.0, SEM = 2.2) and O-P- (mean = 85.0, SEM = 3.0), the two conditions that did not show significant difference between one another.

The Rhyming by Conflict interaction was also significant for response times ($F(1,45) = 6.69$, $p = 0.01$, partial $\eta^2 = 0.13$), and it was due to a significant difference between the conflicting and non-conflicting conditions among the non-rhyming pairs ($t(45) = 4.64$, $p < 0.001$; O+P- and O-P-), and nonsignificant differences between conflicting and nonconflicting conditions among the rhyming pairs ($t(49) = -1.64$; $p = 0.11$; O+P+ and O-P+).

Behavioral results: Performance on the multiplication verification task

We calculated accuracy (i.e., percentage of correctly solved trials) and means of response times (i.e., for correctly solved trials) separately for small and large multiplication problems, for every participant. We then calculated a repeated measures ANOVA including Problem size (i.e., small, large) as within-subjects factor for accuracy and for RTs. For accuracy, we found a main effect of Problem Size ($F(1,49) = 109.39$, $p < 0.001$, partial $\eta^2 = 0.69$), with small problems having higher accuracy than large problems. The main effect of Problem Size was also significant for response times ($F(1,49) = 109.39$, $p < 0.001$, partial $\eta^2 = 0.69$), showing that small problems were solved faster than large problems. More specific information about performance on the multiplication task for the whole sample is given in Table 1.

Table 3 Regression results with behavioral measures

Model	Predictor	Stand. β	t	R^2	ΔR^2	F	ΔF
1	Multiplication skill T1	0.407	3.089**	0.166	0.166	9.539**	9.539
2	Multiplication skill T1	0.314	2.631*	0.359	0.193	13.171**	14.182
	Full IQ T1	0.449	3.766**				
3	Multiplication skill T1	0.307	2.397*	0.360	0.000	8.606**	0.024
	Full IQ T1	0.447	3.658**				
	Math attitudes T1	.020	0.156				

Results for the three regression models used to predict multiplication skill at T2 from multiplication skill at T1 (model 1), multiplication skill, and full IQ at T1 (model 2), and multiplication skill, full IQ, and math attitudes at T1 (model 3).

T1: Time 1; T2: Time 2; * Significant at $p < 0.05$; ** Significant at $p < 0.005$. Dependent measure: Multiplication skill at T2. ΔR^2 : change in R^2 . ΔF : Change in F. No multicollinearity was found in these analyses, with the VIF values ranging from 1.04 to 1.19.

Behavioral results: the role of math attitudes in predicting improvement in multiplication skill

We ran a hierarchical regression analysis to test whether math attitudes at T1 predicted multiplication skill at T2, after accounting for initial levels of multiplication skill (i.e., T1) and full IQ at T1. Multiplication skill at T1 was the only predictor in Model 1. In Model 2, both multiplication skill and full IQ at T1 were included as predictors. Model 3 included multiplication skill, full IQ, and math attitudes, all at T1, as predictors. The dependent measure in all three models was multiplication skill at T2.⁴

As shown in Table 3, results showed that multiplication performance at T2 was significantly explained by multiplication performance at T1 and by full IQ, but that math attitudes at T1 was not a significant predictor over and above these two cognitive factors.

fMRI results

Math attitudes by improvement in multiplication skill interaction in the left IFG

The regression analysis revealed a significant cluster (Fig. 5A; cluster in blue) showing activation for the interaction between math attitudes and improvement in multiplication skill while children solved small multiplication problems inside the scanner. More information about this cluster is given in Table 5. As shown in Fig. 5D, this cluster overlapped the cluster

showing a math attitude by math skill interaction at T1, as found by Demir-Lira and colleagues (2019; cluster in cyan⁵), the peaks from the two clusters being within 24 mm Euclidean distance from each other. Follow-up analyses were performed to understand the interaction of math attitudes by improvement in multiplication skill (see Sections 4.2. and 4.3).

No cluster reached significance for the main effect of math attitudes or the main effect of improvement in multiplication skill. No significant result was obtained for the analysis looking at brain activation elicited when children solved large multiplication problems inside the scanner.

Exploring the interaction: no improvement-related effects in left IFG for math attitude groups

This follow-up effect analysis consisted of running a cluster-wise regression analysis looking for clusters showing an association with improvement in multiplication skill separately for negative and positive math attitudes groups. Full IQ and accuracy in the multiplication task solved inside the scanner also were included as covariates of no interest. To this aim, two groups were formed based on the median-split of raw scores on the math attitudes subtest of the Test of Mathematical Abilities (Brown et al., 2012) at T1. For more information about this test, see Section 2.2.1. The negative math attitudes group comprised 25 children scoring below the median and whose standard scores corresponded to “below average” math attitudes. The positive math attitudes group comprised 25 children scoring above the median and whose standard scores corresponded to “above average” math attitudes. Math attitudes groups did not differ in age at T1 ($t(48) = -0.50, p = 0.62$), age at T2 ($t(48) = -0.40, p = 0.69$), time between sessions ($t(48) = 0.40, p = 0.69$), multiplication skill at T1 ($t(48) = -1.60, p = 0.12$), multiplication skill at T2 ($t(48) = -0.37, p = 0.71$), improvement in multiplication skill ($t(48) = 0.28, p = 0.78$), full IQ at T1 ($t(48) = -1.36, p = 0.18$), sex distribution ($\chi^2 = 0, p = 1$), level of education of the mother ($\chi^2 = 2.73, p =$

⁴ The same results were found if residuals of the change in multiplication skill over time after accounting for T1 was used as the dependent measure, with full IQ significantly predicting improvement in multiplication skill whereas math attitudes at T1 did not.

⁵ Note the difference in brain orientation in our figure and the one in Supplementary Figure 2 in Demir-Lira et al., (2019) paper.

Table 4 Information on the groups created to study the interaction

	Math attitudes groups (median-split)		Improvement groups (median-split)	
	Negative attitudes (n = 25)	Positive attitudes (n = 25)	Improvers (n = 25)	Nonimprovers (n = 25)
Age at T1 session (yr)	11.0 (1.5)	11.2 (1.4)	10.8 (1.3)	11.5 (1.6)
Age at T2 session (yr)	13.2 (1.7)	13.4 (1.5)	12.9 (1.4)	13.7 (1.7)
Time between sessions (yr)	2.2 (0.3)	2.2 (0.2)	2.2 (0.2)	2.2 (0.3)
Female/male ratio	13/12	13/12	12/13	14/11
Math attitudes at T1 ^a	35.8 (4.7)	49.6 (4.3)	43.5 (9.4)	42.0 (7)
Multiplication skill T1 ^a	10.6 (6)	13.0 (4.9)	12.0 (6.1)	11.6 (5.1)
Multiplication skill T2 ^a	13.7 (6)	14.3 (5.5)	18.2 (3.7)	9.8 (4)
Improvement T2-T1 ^a	0.2 (5.3)	-0.2 (5.2)	4.1 (2.7)	-4.1 (3.5)
Reading skill at T1 ^b	107.4 (9.9)	106.0 (12.7)	106.8 (10.3)	103.6 (11.2)
Full IQ at T1 ^b	109.5 (13.8)	115.1 (15.3)	118.5 (14)	106.1 (12.8)
% Accuracy: small	87.1 (14.2)	91.8 (11.1)	93.2 (8.8)*	85.7 (15.2)*
% Accuracy: large	67.0 (17.2)	69.0 (.02)	74.3 (17.8)**	61.7 (14)**
RTs (sec): small	1.21 (.37)	1.11 (.35)	1.19 (.34)	1.14 (0.38)
RTs (sec): large	1.37 (.37)	1.36 (.41)	1.44 (.38)	1.29 (0.39)

Means (standard deviation in parenthesis) for age, time between sessions, and scores on standardized tests of math attitudes, multiplication skill, improvement, reading skill, intelligence (IQ), and accuracy (percentage) and response times (RTs; sec) for small and large multiplication problems solved inside the scanner, for the math attitudes groups (i.e., negative attitudes; positive attitudes; n = 25 each) and for the improvement groups (i.e., improvers; nonimprovers; n = 25 each) created to study the math attitudes by improvement interaction.

^aRaw scores; ^bStandard scores. T1: Time 1; T2: Time 2. See section 2.2. for a description of the standardized tests included in this table. * Significant differences between improvement groups for small problems ($t(48) = -2.13, p = 0.04$). ** Significant differences between improvement groups for large problems ($t(48) = -2.76, p = 0.008$).

0.60), or level of education of the father ($\chi^2 = 2.42, p = 0.66$). As expected, groups significantly differed in math attitudes ($t(48) = -10.83, p < 0.001$). Group differences in math attitudes at T1 are shown in Fig. 4A. More detailed information about these two groups is given in Table 4. As shown in Table 4, these groups did not differ in their performance on the multiplication task solved inside the scanner.⁶

We then carried out two separate cluster-wise regression analyses, for the negative math attitudes (n = 25) and the positive math attitudes (n = 25) groups, looking for clusters within the ROI that showed an association with improvement in multiplication skill over time while children solved small multiplication problems inside the scanner. No cluster reached significance for this analysis.

⁶ We performed a repeated measures ANOVA including Problem size (i.e., small, large) as within-subjects factor and Math attitudes groups (i.e., low, high) as the between-subjects factor, for accuracy and for response times. As for accuracy, we found a main effect of Problem size ($F(1,48) = 108.10, p < 0.001, \text{partial } \eta^2 = 0.69$), but no interaction of Problem size by Math attitudes groups ($F(1,48) = 0.42, p = 0.52, \text{partial } \eta^2 = 0.009$). The main effect of Math attitudes group did not reach significance ($F(1,48) = 0.82, p = 0.37, \text{partial } \eta^2 = 0.02$). As indicated in Table 4, the main effect of Problem Size was due to higher accuracy for small problems as compared to large ones. As for response times, we found a main effect of Problem Size ($F(1,48) = 30, p < 0.001, \text{partial } \eta^2 = 0.38$), but no interaction of Problem Size by Math attitudes groups ($F(1,48) = 1.38, p = 0.25, \text{partial } \eta^2 = 0.03$). The main effect of Math attitudes group did not reach significance ($F(1,48) = 0.29, p = 0.59, \text{partial } \eta^2 = 0.006$). As shown in Table 4, the main effect of Problem Size was due to faster response times for small problems than for large ones.

Exploring the interaction: math attitudes effects in left IFG only for improvers

This follow-up effect analysis consisted of running a cluster-wise regression analysis looking for clusters showing an association with math attitudes separately for improvers and non-improvers groups. Full IQ and accuracy in the multiplication task solved inside the scanner were also included as covariates of no interest. Two groups were formed based on the median-split of the time difference (T2-T1) in raw scores on the multiplication subtests of the Comprehensive Mathematical Abilities Test (CMAT; Hresko et al., 2003), after accounting for initial differences in the test (i.e. residuals after accounting for T1). For more information about this test see section 2.2.2. The nonimprovers group comprised 25 children showing less improvement in multiplication skill over time and the improvers group comprised 25 children showing greater improvement in multiplication skill over time. Groups did not differ in age at T1 ($t(48) = 1.69, p = 0.10$), age at T2 ($t(48) = 1.68, p = 0.10$), time between sessions ($t(48) = 0.58, p = 0.56$), math attitudes at T1 ($t(48) = -0.65, p = 0.52$), multiplication skill at T1 ($t(48) = -0.23, p = 0.82$), sex distribution ($\chi^2 = 0.32, p = 0.78$), level of education of the mother ($\chi^2 = 3.43, p = 0.49$) or level of education of the father ($\chi^2 = 2.76, p = 0.60$). As expected, groups differed in multiplication skill at T2 ($t(48) = -7.70, p < 0.001$) and improvement in multiplication skill ($t(48) = -9.21, p < 0.001$). Groups also differed in full IQ

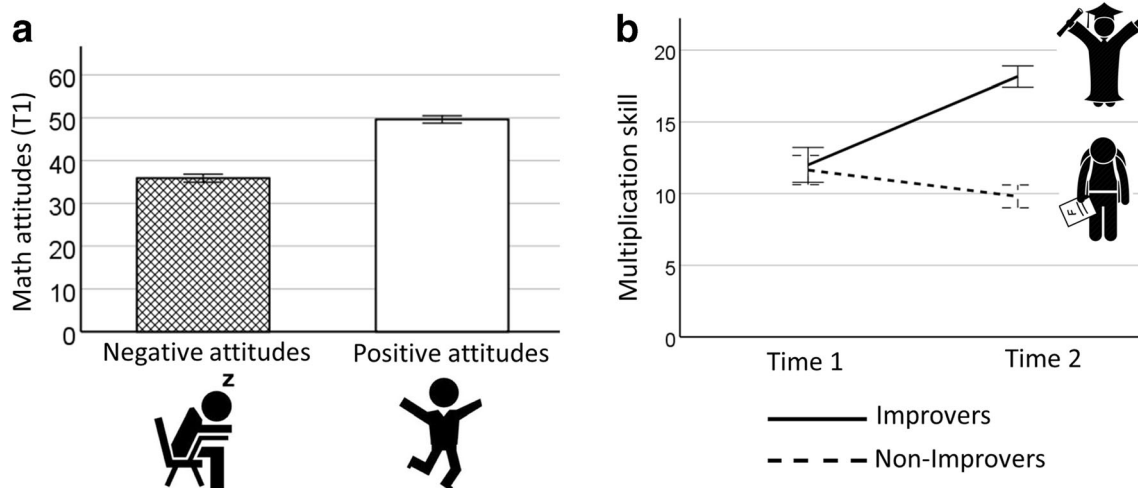


Fig. 4 Math attitudes and improvement groups. **(A)** Illustration of the math attitudes raw scores at T1 for the math attitudes subtest of the Test of Mathematical Abilities (Brown, Cronin, & Bryant, 2012), separately for the groups of children with negative and positive math attitudes (created based on median-split). **(B)** Illustration of the changes over

at T1 ($t(48) = -3.26, p = 0.002$), so this variable was controlled for in the cluster-wise regression analysis (see Section 2.7.3.). Changes in multiplication skill over time for improvers and non-improvers are shown in Fig. 4B. More detailed information about these two groups is given in Table 4. As shown in Table 4, these groups differed on accuracy in the task solved inside the scanner.⁷

We then performed two separate cluster-wise regression analyses, for improvers ($n = 25$) and nonimprovers ($n = 25$) groups, looking for clusters within the ROI that were associated with math attitudes at T1 while children solved small multiplication problems inside the scanner. This analysis revealed a significant cluster showing an association with math attitudes only for the improvers group. This cluster is shown in green in Fig. 5B. As shown in Fig. 5D, this follow-up cluster (i.e., in green) overlapped the cluster showing the interaction

time in the raw scores for the Comprehensive Mathematical Abilities Test (Hresko, Schlieve, Herron, Swain, & Sherbenou, 2003) used to measure multiplication skill, separately for improvers and non-improvers children (created based on median-split). Error bars show standard error of the mean.

of math attitudes by improvement in multiplication skill (i.e., in blue), the peaks from the two clusters being just 2-mm Euclidean distance from each other. This follow-up cluster also overlapped the cluster that Demir-Lira et al., (2019) found to be significant for the interaction of math attitudes by math skill at T1 (i.e., in cyan in Fig. 5), with the peaks from the two clusters being within 24-mm Euclidean distance from one another. More specific information about this cluster is provided in Table 5.

For illustration purposes, we then extracted brain activation from this follow-up cluster from children showing improvement and plotted it against math attitudes scores. As shown in Fig. 6's scatterplot for the improvers groups, the more positive

⁷ We performed a repeated measures ANOVA including Problem size (i.e. small, large) as within-subjects factor and Improvement groups (i.e., improvers; nonimprovers) as the between-subjects factor, for accuracy and for response times. As for accuracy, we found a main effect of Problem Size ($F(1,48) = 110.60, p < 0.001, \text{partial } \eta^2 = 0.70$), showing higher accuracy for small problems than for large ones, but no interaction of Problem Size by Improvement groups ($F(1,48) = 1.54, p = 0.22, \text{partial } \eta^2 = 0.03$). The main effect of Improvement groups was significant ($F(1,48) = 8.17, p = 0.006, \text{partial } \eta^2 = 0.15$). As indicated in Table 4, the main effect of Improvement groups was due to improvers showing higher accuracy than nonimprovers, both for small and large multiplication problems. For this reason, accuracy in small and large multiplication problems at T1 were included as covariates of no interest (i.e. control variables) in the cluster-wise regression analyses described in section 2.7.3. As for response times, we found a main effect of Problem Size ($F(1,48) = 29.79, p < 0.001, \text{partial } \eta^2 = 0.38$), but no interaction of Problem Size by Improvement groups ($F(1,48) = 1.46, p = 0.23, \text{partial } \eta^2 = 0.03$). The main effect of Improvement groups did not reach significance ($F(1,48) = 1.03, p = 0.31, \text{partial } \eta^2 = 0.02$). The main effect of Problem Size was due to faster response times to small problems compared with large ones.

Table 5 Interaction and follow-up cluster information

K	aal	~BA	MNI coordinate			Z value
			X	Y	Z	
Math attitudes by improvement in multiplication skill interaction						
58	IFG Triangularis; IFG Operularis	45, 44	-40	12	28	3.27
			-32	6	30	3.24
Follow-up: Main effect of math attitudes for improvers						
57	IFG Triangularis; IFG Operularis	45, 44	-42	12	28	3.74
			-50	8	28	2.99

Cluster size (K), corresponding region based on anatomical automatic labeling (aal), approximate Brodmann areas (~BA), MNI coordinates of the peaks, and Z-values for the cluster showing the interaction of math attitudes by improvement in multiplication skill ($K = 58$) and for the cluster showing the main effect of math attitudes only for the improvement group ($K = 57$)

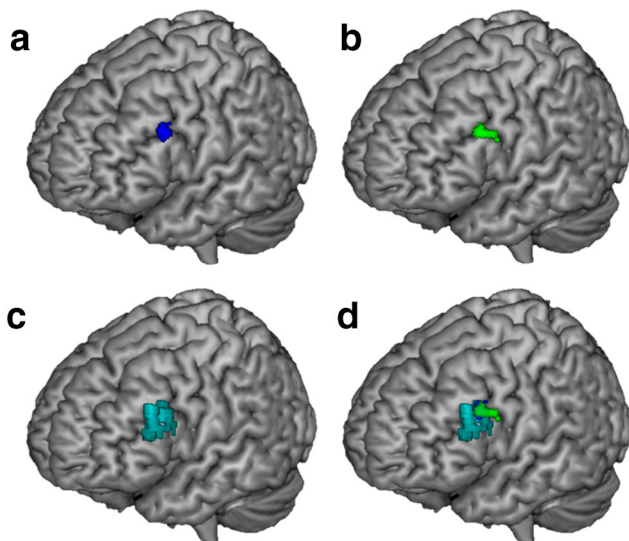


Fig. 5 Clusters showing significance for the interaction and the follow-up analyses. **(A)** Cluster showing significance for the interaction between math attitudes and improvement in multiplication skill over time (i.e., in blue; $K = 58$). **(B)** Cluster found for the follow-up analysis showing math attitudes-related activation only for improvers (i.e., in green; $K = 57$). **(C)** Cluster found by Demir-Lira et al. (2019) showing an interaction between math attitudes and math skill at T1 (i.e., in cyan; $K = 127$). **(D)** Overlap of the three clusters: Cluster showing the follow-up activation for improvers (i.e., in green; $K = 57$), overlaid on the cluster showing the interaction of math attitudes by improvement in multiplication skill (i.e., in blue; $K = 58$), and on the Demir-Lira et al. (2019) cluster showing the interaction of math attitudes by math skill at T1 (i.e., in cyan; $K = 124$).

their math attitudes, the more they activated this cluster in left IFG to solve small multiplication problems inside the scanner.

No cluster reached significance for the nonimprovers group.

Nonimprovers with positive math attitudes decreased math attitudes over time

As indicated above, the follow-up analysis showed no significant effect of attitudes for the nonimprovers. We were interested in the subgroup of children showing initial positive levels of math attitudes but not showing improvement in multiplication skill over time.⁸ We considered that this lack of improvement might be because, despite their positive math attitudes at T1, these children may have decreased their math attitudes over time. To explore this, we ran a paired-sample T-test comparing math attitudes scores at T1 and T2 for the subgroup of nonimprovers with positive math attitudes at T1. We found a significant change in math attitudes over time ($t(12) = 3.60, p = 0.004$) for this group, with children decreasing math attitudes scores from T1 (mean = 47.15; SD = 3.10) to T2 (mean = 41; SD = 5.58).

Whole-brain results

Four clusters reached significance for the regression analysis with small multiplication problems at the whole-brain level. Two of these clusters were located in the lingual gyrus ($k = 314$; peak coordinates: $x = -12, y = -64, z = -6$; $Z = 3.80$ and $k = 393$; peak coordinates: $x = 30, y = -76, z = -16$; $Z = 3.66$). The third cluster ($k = 216$) involved the sub-gyral (peak coordinates: $x = 30, y = -80, z = 20$; $Z = 3.61$) and the cuneus (peak coordinates: $x = 26, y = -82, z = 28$; $Z = 3.45$). The fourth cluster ($k = 237$) included the extra nuclear (peak coordinates: $x = 6, y = -6, z = 2$; $Z = 3.51$) and the thalamus (peak coordinates: $x = -4, y = -18, z = 12$; $Z = 3.40$). No cluster reached significance for the regression analysis with large multiplication problems.

Discussion

To succeed in our technological society, students need to develop their capabilities in science, engineering, and mathematics to levels beyond what was considered acceptable in the past (National Science Board, 2007). Because math achievement is the result of both cognitive (Geary, 2011) and non-cognitive factors (McLeod, 1992), both must be addressed when considering how to improve children's math learning and education.

Several longitudinal studies have shown that math attitudes can predict math improvement, above and beyond initial levels of math skill (Hemmings et al., 2011; Wigfield & Eccles, 2000). The investment of effort, usually defined as the mobilization of resources that are associated with an action (Hepler & Albarracin, 2014), is considered one of the possible mechanisms by which math attitudes impact math performance. Hemmings and Kay (2010) found a positive

⁸ These subgroups resulted in 12 participants being classified as nonimprovers with negative math attitudes, 13 participants as nonimprovers with positive math attitudes, 13 participants as improvers with negative math attitudes, and 12 participants as improvers with positive math attitudes. The four subgroups did not differ in age at T1, age at T2, time between sessions, reading skill, or multiplication skill at T1 (all p values of Bonferroni-corrected post hoc tests being above 0.42). In addition, the subgroups did not differ in sex distribution ($X^2 = 0.41, p = 0.94$). The two negative attitudes groups did not differ from one another in attitudes ($p = 1.0$), but the two positive attitudes groups did differ ($p = 0.20$), with the positive attitudes improver's group having more positive attitudes (mean = 52.3, SD = 3.8) than their nonimprovers counterparts (mean = 47.2, SD = 3.1). As expected, the subgroups also differed in improvement, with the two improvement subgroups differing from the two non-improvement ones (all p values < 0.001). The two subgroups showing improvement (i.e. with positive and negative attitudes) did not differ from each other and the two subgroups showing non-improvement did not differ from one another (all p values = 1.0). The two subgroups differed from each other in full IQ ($p = 0.007$). The improvers with positive math attitudes showed higher IQ (mean = 121.1, SD = 15.0) than the group of nonimprovers with negative attitudes (mean = 102.3, SD = 10.8).

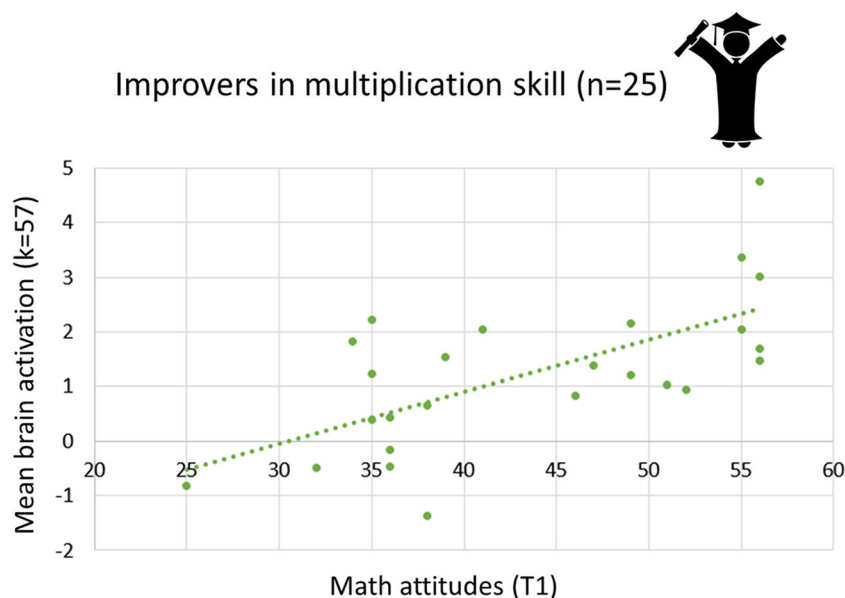


Fig. 6 Association between math attitudes and brain activation for the improvers group shown for illustration purposes. Scatterplot showing the positive correlation between math attitudes, in the X-axis, and the average

correlation ($r = 0.55$) between effort investment and math attitudes in 16-year-old students. In a study with college students, Cole et al. (2008) found that attitudinal aspects, such as interest and perceived usefulness of a test, led to better math achievement through more effort invested in the test, as measured by self-report measures. Trautwein et al. (2009) studied eighth graders and found that math attitudes at T1 predicted math achievement at T2 through more academic effort, as measured with items referring to working hard to do better in math. Balfanz and Byrnes (2006) found that effort was an important factor explaining the reduction of the math achievement gap in high poverty middle schools, with students that indicated that they were working “as hard as I can in math class” having a 19% greater probability of catching up than those who said they were “not working hard at all.” Overall, this evidence suggests that effort investment might be the mechanisms explaining the predictive relation between math attitudes and math achievement.

Despite the extensive literature on the relationship between math attitudes and math performance using behavioral measures, very little is known about the neurocognitive mechanisms explaining this association. Only two studies have investigated the brain correlates of their association (Chen et al., 2018) and interaction (Demir-Lira et al., 2019) using fMRI. The correlational study by Chen et al. (2018) found that the association between math attitudes and math performance was mediated by increased use of a retrieval strategy and greater hippocampal activation. In another study, Demir-Lira et al. (2019) found an interaction between math attitudes and math skill while children solved a multiplication task. The interaction was due to low skilled children with positive math attitudes activating the left IFG to a greater degree as compared with the other groups.

brain activation extracted from the cluster showing significance in the follow-up analysis (i.e., $K = 57$; in green in Fig. 5B), in the Y-axis. Error bars show standard error of the mean.

Greater activation in the left IFG was interpreted as reflecting the investment of more effort in retrieving multiplication facts from memory. However, because of the cross-sectional nature of these studies, neither was able to establish the directionality of the effects between math attitudes and math achievement. In this study, we examined predictive relations between math attitudes and math improvement using longitudinal data, which move us one step closer to establishing the directionality of the effects between these variables.

No longitudinal study has yet investigated the underlying neurocognitive mechanisms by which positive math attitudes lead to improvements in math skill over time, which was the objective of the current study. While Demir-Lira et al. (2019) showed that positive math attitudes were associated with lower skill children’s activation in the left IFG, the current study was designed to answer whether this same brain mechanism could explain children’s improvement in multiplication skill over time, regardless of initial levels of performance. We first identified a region of interest in the left IFG implicated in verbal processing using a rhyming judgment task. We then performed a regression analysis within this ROI during multiplication problem-solving. Our analysis showed an interaction of math attitudes by improvement in multiplication skill, which overlapped the IFG cluster found by Demir-Lira et al. (2019). Follow-up analyses showed that this interaction was due to an association between math attitudes and left IFG activation only for improvers, with more positive math attitudes related to greater left IFG activation. In other words, the interaction was due to improvers with positive math attitudes showing greater activation in the left IFG compared with those with negative math attitudes. This result confirms our hypothesis that enhanced left IFG activation is not only the neural

correlate of positive attitudes in children with lower skill (Demir-Lira et al., 2019), but also the mechanism by which positive math attitudes explain improvement in multiplication skill over time.

As with Demir-Lira et al. (2019), we interpreted the greater left IFG activation as reflecting effort invested in problem-solving. Neuroimaging studies have shown that children show reduced activation of this region with more years of math instruction (Prado et al., 2014) and that young adults show less IFG activation compared with children (Rivera et al., 2005). These results suggest less effortful retrieval of arithmetic facts as math expertise develops. However, longitudinal studies have shown that brain activation of the left IFG during verbal processing was a significant predictor of gains in multiplication skill over time, with improvement being associated with less initial activation in this region (Suárez-Pellicioni, Fuchs and Booth, 2019).⁹ Does our current finding, showing greater activation in IFG at T1 predicting larger longitudinal gains, contradict this evidence? While some studies have focused on age (Rivera et al., 2005), years of math instruction (Prado et al., 2014), or improvement (Suárez-Pellicioni, Fuchs, & Booth, 2019) as their variables of interest, in the present study we focused on the interaction between math attitudes and multiplication improvement. Other possible explanations for the lack of age-related effects in our findings could be that: 1) There was no deterioration in math attitudes with age (Ma & Kishor, 1997) in our sample ($r = 0.10, p = 0.51$). This could have contributed to not finding an age-related effect when studying the interaction between math attitudes and improvement. 2) There was a significant correlation between age at T1 and multiplication skill at T1 ($r = 0.39, p = 0.006$). This was due to older children showing higher achievement at T1 compared with younger ones. It is possible that by accounting for initial levels of skill when calculating our measure of improvement (i.e., residuals), we may have in part accounted for these age-related effects. 3) There was a significant negative correlation between age at T1 and improvement in multiplication skill ($r = -0.30, p = 0.04$). This was due to younger children showing more improvement than older ones. This is potentially due to younger children showing lower performance at T1, so they had more “room” for improvement over time. We included the main effect of improvement as a covariate of no interest in

the regression analyses, so we may have in part controlled for these age-related effects. These methodological aspects may have contributed to the fact that we did not find age-related IFG effects as reported by other studies (Prado et al., 2014; Rivera et al., 2005).

We interpreted IFG activation as effort invested in the task,¹⁰ in line with previous math research showing greater IFG activation in situations that require more effortful processing, such as when adults solved untrained multiplication problems as compared to trained ones (Delazer et al., 2003), when children solved two-digit multiplications as compared to single-digit ones (Soltanlou et al., 2017), or when nonimproving children relied on less efficient strategies to solve multiplication problems inside the scanner (Suárez-Pellicioni et al., 2018). In addition to the literature on math cognition, many studies have shown that the left IFG is critical for tasks with effortful demands (Massar et al., 2015). A meta-analysis of 53 fMRI studies investigating auditory perception found that the left IFG was associated with effortful listening under different speech manipulations (Alain et al., 2018). In a study looking at verbal learning of easy and hard words after normal sleep and following 36 hours of sleep deprivation, the left IFG was one of the regions associated with better free recall performance of hard words after sleep deprivation, the condition reported as requiring greater effort (Drummond et al., 2005). The IFG was activated while participants rated the level of difficulty of a memory task with different load conditions, with difficulty level correlating with self-reported ratings of mental effort invested in the task (Otto et al., 2014). Aben et al. (2020) asked participants to detect faces in noisy images, with effort demands being manipulated by adding noise to them. They found that the IFG was activated when comparing the high effort versus low effort conditions. Finally, Poldrack et al. (2001) presented participants with a verification task including speech with different degrees of temporal compression. They found that the left IFG showed increased activation with increasing speech compression, suggesting more effortful processing.

The above-mentioned literature shows greater frontal activation for more challenging tasks/conditions. Why did we find the effects for small, but not large, multiplication problems solved inside the scanner? This could be due to the fact that we aimed to study improvement regardless of initial levels of performance. After accounting for skill-related differences, the effect of math attitudes was not found for large multiplication problems because skill plays a more important role in difficult tasks. In other words, our results suggest that attitudes-related effects are more evident in conditions in which skill plays a less important role and that for more challenging conditions, in which skill level may play a more dominant role, the effects of attitudes may be harder

⁹ Note that Suárez-Pellicioni, Fuchs, and Booth (2019) focused on the role of IFG in phonological processing in predicting math skill and found that less activation in this region predicted math gains only for younger children (ages 8–11 at T1).

¹⁰ It also is possible that the enhanced IFG activation for the improvers could be instead reflecting motivation in participants with positive math attitudes. Future studies should directly measure effort and motivational aspects in order to provide stronger conclusions regarding the underlying mechanisms explaining the effects of positive math attitudes on math improvement.

to detect. In contrast to the scanner task, the test used to measure improvement included a wide range of difficulty. This test was also untimed, because we aimed to measure what children were *willing* to do when there were no time constraints. We expected untimed tests to be more sensitive to perseverance- and effort-related effects, whereas timed tests may be more vulnerable to the effect of anxiety (Beilock & Carr, 2005). Math anxiety, which was not measured in this study, shows a detrimental effect on math performance (Ashcraft, 2002; Suárez-Pellicioni et al., 2016). Studies have reported negative correlations between math anxiety and math attitudes as a whole (ranging from $r = -0.47$ to $r = -0.88$; Hembree, 1990) and with its interests and perceptions components ($r = -0.45$ to $r = -0.70$; Malanchini et al., 2020), indicating that children with positive math attitudes usually score low in math anxiety. Considering the negative correlation between these variables and given the untimed nature of the test used, we think it is unlikely that the enhanced IFG activation we found for the improvers with positive math attitudes is due to anxiety.

The longitudinal nature of our study allowed us to provide evidence that better speaks the directionality of the effects between math attitudes (T1) and improvement (T2-T1), suggesting that engaging left IFG *early on* by those children with positive math attitudes drives gains in multiplication skill over time. Our study suggests that investing effort in retrieving multiplication facts from memory, as shown by enhanced left IFG activation, is a possible mechanism by which children with positive math attitudes improve in an untimed standardized test of multiplication skill. However, not all children with positive math attitudes at T1 showed improvement in multiplication skill. Although we found no neural correlate of lack of improvement for those children with positive attitudes, our analysis revealed that attitudes became more negative over time for this group. It is possible that attitudes becoming more negative impacts math performance, and explains our finding of lack of improvement. This interpretation is consistent with evidence showing that math attitudes may play a causal role in performance (Ma & Kishor, 1997). Alternatively, it could be the case that lack of improvement led to math attitudes becoming more negative over time. This interpretation is consistent with evidence suggesting that math achievement causes changes in math attitudes (Ma & Xu, 2004). A third possibility is that math attitudes and math achievement show a reciprocal relationship, affecting one another (Ganley & Lubienski, 2016; Pinxten et al., 2014).

Despite our finding that brain measures predict outcomes, our regression analysis with behavioral measures showed that initial levels of math attitudes did not predict multiplication skill at T2, after accounting for multiplication skill and intelligence at T1. Our lack of behavioral prediction supports previous behavioral results suggesting that the effect of math

attitudes were minimal (Hemmings & Kay, 2010; Reynolds & Walberg, 1992). Our findings go in line with previous fMRI evidence showing that brain measures were significant predictors of skill gains in math (Demir-lira et al., 2016) and in other domains, such as reading (Hoeft et al., 2011), whereas behavioral measures were not. Neural measures may be more sensitive in predicting outcomes than self-reported measures of math attitudes. However, our questionnaire of math attitudes asked about liking, engagement/avoidance, self-perceptions of math skill, and perceived usefulness of math in general. Some researchers have argued for more precise measures of attitudes toward different math domains, such as arithmetic, geometry, and problem-solving (Aiken, 1970b). Studies have shown, for example, negative attitudes toward fractions but positive attitudes toward whole numbers both for children and adults, regardless of math achievement levels (Sidney et al., 2019). This specificity of our brain measure compared with the more broadly defined measure of math attitudes obtained from the standardized test may explain the discrepancy between behavioral and fMRI results. We predict that left IFG activation is a good candidate for predicting gains in operations that are solved by retrieving facts from memory, but not for subtraction operations that rely on quantity processing in parietal cortex (Prado et al., 2011; Prado et al., 2014; Suárez-Pellicioni et al., 2020).

Our study proposes that greater left IFG activation is associated with the investment of effort and represents the neurocognitive mechanisms by which positive math attitudes lead to improvement in multiplication skill over time. Studies have shown that students are more likely to work harder if they attribute success or failure to effort rather than to ability (Schunk, 1983). Cultural differences have been found in success attributions. Students in the United States tend to believe that ability is the main cause of achievement, whereas students in Japan tend to attribute success to effort (Holloway, 1988). Changing students' success attributions, perhaps emphasizing Thomas Edison's idea that "Genius is 1% inspiration and 99% perspiration," could have a positive impact on children's willingness to invest effort in math. Teachers and parents could promote the importance of effort and practice for mastering math instead of emphasizing ability and should pay attention to how they praise children. Studies have shown that praising effort led to greater task persistence, more task enjoyment, and better task performance compared with praising intelligence (Mueller & Dweck, 1998).

Because of the malleable nature of math attitudes, teachers and parents can help children make the most out of their potential by targeting their math attitudes. Simple school interventions, such as working in cooperative groups where students help each other, has been shown to improve math attitudes (Leikin & Zaslavsky, 1997). These interventions might bring greater benefits if they take place early in elementary school, certainly, before

children have the chance to avoid mathematics in their high school curricular choices (Singh et al., 2002). Studies also have shown that interventions to improve both math skill and student's attitudes are more effective than those focusing only on math skill (Marsh & Craven, 2006; Seaton et al., 2014). Only the study of both the cognitive and affective components involved in math learning will provide a comprehensive explanation of why and how some children excel in math while others are left behind, which will help teachers and parents in their work of preparing children to become productive members of society.

Acknowledgments This research was supported by HD059177 from the National Institute of Child Health and Human Development to JRB. The authors thank John Binzak and Rachna Mutreja for their help with data collection. They also thank all participating children and their parents.

References

- Aben, B., Calderon, C. B., van den Bussche, E., & Verguts, T. (2020). Cognitive effort modulates connectivity between dorsal anterior cingulate cortex and task-relevant cortical areas. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.2948-19.2020>
- Ai, X. (2002). Gender Differences in Growth in Mathematics Achievement: Three-Level Longitudinal and Multilevel Analyses of Individual, Home, and School Influences. *Mathematical Thinking and Learning*. https://doi.org/10.1207/s15327833mtl0401_1
- Aiken, L. R. (1970a). Attitudes toward mathematics. *Review of Educational Research*. <https://doi.org/10.3102/00346543040004551>
- Aiken, L. R. (1970b). Nonintellective variables and mathematics achievement: Directions for research. *Journal of School Psychology* [https://doi.org/10.1016/0022-4405\(70\)90052-X](https://doi.org/10.1016/0022-4405(70)90052-X)
- Alain, C., Du, Y., Bernstein, L. J., Barten, T., & Banai, K. (2018). Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Human Brain Mapping*. <https://doi.org/10.1002/hbm.24031>
- Andin, J., Fransson, P., Rönnerberg, J., & Rudner, M. (2015). Phonology and arithmetic in the language – calculation network. *Brain and Language*, 143, 97–105. <https://doi.org/10.1016/j.bandl.2015.02.004>
- Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science* <https://doi.org/10.1111/1467-8721.00196>
- Aubrey, C., Godfrey, R., & Dahl, S. (2006). Early mathematics development and later achievement: Further evidence. *Mathematics Education Research Journal* <https://doi.org/10.1007/BF03217428>
- Aunola, K., Leskinen, E., & Nurmi, J. E. (2006). Developmental dynamics between mathematical performance, task motivation, and teachers' goals during the transition to primary school. *British Journal of Educational Psychology* <https://doi.org/10.1348/000709905X51608>
- Balfanz, R., & Byrnes, V. (2006). Closing the Mathematics Achievement Gap in High-Poverty Middle Schools: Enablers and Constraints. *Journal of Education for Students Placed at Risk (JESPAR)* https://doi.org/10.1207/s15327671espr1102_2
- Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and “Choking under pressure” in math. *Psychological Science* <https://doi.org/10.1111/j.0956-7976.2005.00789.x>
- Berteletti, I., Man, G., & Booth, J. R. (2014). How number line estimation skills relate to neural activations in single digit subtraction problems. *NeuroImage*, 107C, 198–206. <https://doi.org/10.1016/j.neuroimage.2014.12.011>
- Berteletti, I., & Booth, J. R. (2015). Perceiving fingers in single-digit arithmetic problems. *Frontiers in Psychology*, 6(March), 1–10. <https://doi.org/10.3389/fpsyg.2015.00226>
- Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *Elementary School Journal*. <https://doi.org/10.1086/525550>
- Bookheimer, S. (2002). Functional MRI and Language: New Approaches to Understanding the Cortical Organization of Semantic Processing. *Annual Review of Neuroscience*, 25, 151–188. <https://doi.org/10.1146/annurev.neuro.25.112701.142946>
- Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2003). Relation between brain activation and lexical performance. *Human Brain Mapping*, 19(3), 155–169. <https://doi.org/10.1002/hbm.10111>
- Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2004). Development of Brain Mechanisms for Processing Orthographic and Phonologic Representations. *Journal of Cognitive Neuroscience*, 16(7), 1234–1249. <https://doi.org/10.1021/ja064902x.Brilliant>
- Borman, G. D., & Overman, L. T. (2004). Academic resilience in mathematics among poor and minority students. *Elementary School Journal*. <https://doi.org/10.1086/499748>
- Brown, V. L., Cronin, M. E., & Bryant, D. H. (2012). *Test of mathematical abilities*. (PRO-ED).
- Chen, L., Bae, S. R., Battista, C., Qin, S., Chen, T., Evans, T. M., & Menon, V. (2018). Positive Attitude Toward Math Supports Early Academic Success: Behavioral Evidence and Neurocognitive Mechanisms. *Psychological Science*. <https://doi.org/10.1177/0956797617735528>
- Cole, J. S., Bergin, D. A., & Whittaker, T. A. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology* <https://doi.org/10.1016/j.cedpsych.2007.10.002>
- Crombie, G., Sinclair, N., Silverthorn, N., Byrne, B. M., Dubois, D. L., & Trinneer, A. (2005). Predictors of young adolescents' math grades and course enrollment intentions: Gender similarities and differences. *Sex Roles* <https://doi.org/10.1007/s11199-005-2678-1>
- Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic - An fMRI study. *Cognitive Brain Research* <https://doi.org/10.1016/j.cogbrainres.2003.09.005>
- Demir-Lira, E., Suárez-Pellicioni, M., Binzak, J. V., & Booth, J. R. (2019). Attitudes Toward Math Are Differentially Related to the Neural Basis of Multiplication Depending on Math Skill. *Learning Disability Quarterly* <https://doi.org/10.1177/0731948719846608>
- Demir-lira, Ö. E., Prado, J., & Booth, J. R. (2016). Neural Correlates of Math Gains Vary Depending on Parental Socioeconomic Status (SES). *Frontiers in Psychology*, 7(June), 1–12. <https://doi.org/10.3389/fpsyg.2016.00892>
- Demir, Ö. E., Prado, J., & Booth, J. R. (2014). The Differential Role of Verbal and Spatial Working Memory in the Neural Basis of Arithmetic. *Developmental Neuropsychology*, 39(6), 440–458. <https://doi.org/10.1080/87565641.2014.939182>
- Demir, Ö. E., Prado, J., & Booth, J. R. (2015). Parental socioeconomic status and the neural basis of arithmetic: differential relations to verbal and visuo-spatial representations. *Developmental Science*, 5, 1–16. <https://doi.org/10.1111/desc.12268>

- DiPerna, J. C., Lei, P. W., & Reid, E. E. (2007). Kindergarten predictors of mathematical growth in the primary grades: An investigation using the Early Childhood Longitudinal Study - Kindergarten cohort. *Journal of Educational Psychology*. <https://doi.org/10.1037/0022-0663.99.2.369>
- Dowker, A., Cheriton, O., Horton, R., & Mark, W. (2019). Relationships between attitudes and performance in young children's mathematics. *Educational Studies in Mathematics*. <https://doi.org/10.1007/s10649-019-9880-5>
- Drummond, S. P. A., Meloy, M. J., Yanagi, M. A., Orff, H. J., & Brown, G. G. (2005). Compensatory recruitment after sleep deprivation and the relationship with performance. *Psychiatry Research - Neuroimaging* <https://doi.org/10.1016/j.psychres.2005.06.007>
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *PNAS*, 113(33), 7900–7905. <https://doi.org/10.1073/pnas.1612033113>
- Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Report Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area. *Current Biology*, 22(21), 2059–2062. <https://doi.org/10.1016/j.cub.2012.09.011>
- Fisher, P. H., Dobbs-Oates, J., Doctoroff, G. L., & Arnold, D. H. (2012). Early math interest and the development of math skills. *Journal of Educational Psychology* <https://doi.org/10.1037/a0027756>
- Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences* <https://doi.org/10.1016/j.lindif.2016.01.002>
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology* <https://doi.org/10.1037/a0025510>
- Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest Matters: The Importance of Promoting Interest in Education. *Policy Insights from the Behavioral and Brain Sciences*. <https://doi.org/10.1177/2372732216655542>
- Hembree, R. (1990). The Nature, Effects, and Relief of Mathematics Anxiety. *Journal for Research in Mathematics Education*. <https://doi.org/10.2307/749455>
- Hemmings, B., Grootenboer, P., & Kay, R. (2011). Predicting mathematics achievement: The influence of prior achievement and attitudes. *International Journal of Science and Mathematics Education*. <https://doi.org/10.1007/s10763-010-9224-5>
- Hemmings, B., & Kay, R. (2010). Prior achievement, effort, and mathematics attitude as predictors of current achievement. *Australian Educational Researcher*. <https://doi.org/10.1007/BF03216921>
- Hepler, J., & Albarracín, D. (2014). Liking More Means Doing More. *Social Psychology*. <https://doi.org/10.1027/1864-9335/a000198>
- Hoefl, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., Lyytinen, H., Whitfield-Gabrieli, S., Glover, G. H., Reiss, A. L., & Gabrieli, J. D. E. (2011). Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.1008950108>
- Holloway, S. D. (1988). Concepts of Ability and Effort in Japan and the United States. *Review of Educational Research*. <https://doi.org/10.3102/00346543058003327>
- Hresko, W. P., Schlieve, P. L., Herron, S. R., Swain, C., & Sherbenou, R. J. (2003). *Comprehensive mathematical abilities test*. (PRO-ED).
- Leikin, R., & Zaslavsky, O. (1997). Facilitating student interactions in mathematics in a cooperative learning setting. *Journal for Research in Mathematics Education*. <https://doi.org/10.2307/749784>
- Lipnevich, A. A., Preckel, F., & Krumm, S. (2016). Mathematics attitudes and their unique contribution to achievement: Going over and above cognitive ability and personality. *Learning and Individual Differences*. <https://doi.org/10.1016/j.lindif.2015.12.027>
- Ma, X., & Xu, J. (2004). Determining the causal ordering between attitude toward mathematics and achievement in mathematics. *American Journal of Education*. <https://doi.org/10.1086/383074>
- Ma, X., & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for Research in Mathematics Education*. <https://doi.org/10.2307/749662>
- Malanchini, M., Rimfeld, K., Wang, Z., Petrill, S. A., Tucker-Drob, E. M., Plomin, R., & Kovas, Y. (2020). Genetic factors underlie the association between anxiety, attitudes and performance in mathematics. *Translational Psychiatry*. <https://doi.org/10.1038/s41398-020-0711-3>
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on Psychological Science*. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>
- Massar, S. A. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2015.06.080>
- Mazaika, P., Hoefl, F., Glover, G., & Reiss, A. (2009). Methods and software for fMRI analysis for clinical subjects. *The Organization of Human Brain Mapping, 15th Annual Meeting; 2009 Jun 18–23; San Francisco, CA*.
- Mazzocco, M. M. M., Hanich, L. B., & Noeder, M. M. (2012). Primary School Age Students' Spontaneous Comments about Math Reveal Emerging Dispositions Linked to Later Mathematics Achievement. *Child Development Research*. <https://doi.org/10.1155/2012/170310>
- McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In *Handbook of research on mathematics teaching and learning*. New York.
- Merkley, R., Matejko, A. A., & Ansari, D. (2017). Strong causal claims require strong evidence: A commentary on Wang and colleagues. In *Journal of Experimental Child Psychology*. <https://doi.org/10.1016/j.jecp.2016.07.008>
- Mueller, C. M., & Dweck, C. S. (1998). Praise for Intelligence Can Undermine Children's Motivation and Performance. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/0022-3514.75.1.33>
- National Science Board. (2007). National Action Plan for Addressing the Critical Needs of the U.S. Science, Technology, and Mathematics Education System. *October*. <https://doi.org/10.1017/CBO9781107415324.004>
- Neale, D. (1969). The Role of Attitudes in Learning Mathematics. *Arithmetic Teacher*.
- Otto, T., Zijlstra, F. R. H., & Goebel, R. (2014). Neural correlates of mental effort evaluation-involvement of structures related to self-awareness. *Social Cognitive and Affective Neuroscience* <https://doi.org/10.1093/scan/nss136>
- Petersen, J. L., & Hyde, J. S. (2017). Trajectories of self-perceived math ability, utility value and interest across middle school as predictors of high school math performance. *Educational Psychology* <https://doi.org/10.1080/01443410.2015.1076765>
- Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure *British Journal of Educational Psychology* <https://doi.org/10.1111/bjep.12028>
- Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., & Gabrieli, J. D. E. (2001). Relations between the neural bases of dynamic auditory processing and phonological processing: Evidence from fMRI. *Journal of Cognitive Neuroscience* <https://doi.org/10.1162/089892901750363235>
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional Specialization for

- Semantic and Phonological Processing in the Left Inferior Prefrontal Cortex. *NeuroImage*, 35(1), 15–35. <https://doi.org/10.1006/nimg.1999.0441>
- Pollack, C., & Ashby, N. C. (2017). Where Arithmetic and Phonology Meet: The Meta-Analytic Convergence of Arithmetic and Phonological Processing in the Brain. *Developmental Cognitive Neuroscience*, *In press*. <https://doi.org/10.1016/j.dcn.2017.05.003>
- Prado, Jérôme, Mutreja, R., & Booth, J. R. (2014). Developmental dissociation in the neural responses to simple multiplication and subtraction problems. *Developmental Science*, 17(4), 537–552. <https://doi.org/10.1111/desc.12140>
- Prado, Jérôme, Mutreja, R., Zhang, H., Mehta, R., Desroches, A. S., Minas, J. E., & Booth, J. R. (2011). Distinct Representations of Subtraction and Multiplication in the Neural Systems for Numerosity and Language. *Human Brain Mapping*, 32, 1932–1947. <https://doi.org/10.1002/hbm.21159>
- Reynolds, A. J. (1991). The middle schooling process: influences on science and mathematics achievement from the longitudinal study of American youth. *Adolescence*
- Reynolds, Arthur J., & Walberg, H. J. (1992). A Structural Model of High School Mathematics Outcomes. *Journal of Educational Research* <https://doi.org/10.1080/00220671.1992.9944431>
- Rickard, T. C., Romero, S. G., Basso, G., Wharton, C., Flitman, S., & Grafman, J. (2000). The calculating brain : an fMRI study. *Neuropsychologia*, 38, 325–335.
- Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, 15(11), 1779–1790. <https://doi.org/10.1093/cercor/bhi055>
- Schunk, D. H. (1983). Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. *Journal of Educational Psychology* <https://doi.org/10.1037/0022-0663.75.6.848>
- Seaton, M., Parker, P., Marsh, H. W., Craven, R. G., & Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement: Juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology*. <https://doi.org/10.1080/01443410.2013.825232>
- Sidney, P. G., Thompson, C. A., Fitzsimmons, C., & Taber, J. M. (2019). Children’s and Adults’ Math Attitudes Are Differentiated by Number Type. *Journal of Experimental Education* <https://doi.org/10.1080/00220973.2019.1653815>
- Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *Journal of Educational Research* <https://doi.org/10.1080/00220670209596607>
- Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A. C., & Nuerk, H. C. (2017). Increased arithmetic complexity is associated with domain-general but not domain-specific magnitude processing in children: A simultaneous fNIRS-EEG study. *Cognitive, Affective and Behavioral Neuroscience* <https://doi.org/10.3758/s13415-017-0508-x>
- Suárez-Pellicioni, M., Berteletti, I., & Booth, J. (2020). Early Engagement of Parietal Cortex for Subtraction Solving Predicts Longitudinal Gains in Behavioral Fluency in Children. *Frontiers in Human Neuroscience*, 14(163).
- Suárez-Pellicioni, M., & Booth, J. R. (2018). Fluency in symbolic arithmetic refines the approximate number system in parietal cortex. *Human Brain Mapping*, 39, 3956–3971. <https://doi.org/10.1002/hbm.24223>
- Suárez-Pellicioni, M., Fuchs, L., & Booth, J. R. (2019). Temporo-frontal activation during phonological processing predicts gains in arithmetic facts in young children. *Developmental Cognitive Neuroscience*, 40. <https://doi.org/10.1016/j.dcn.2019.100735>
- Suárez-Pellicioni, M., Lytle, M., Younger, J. W., & Booth, J. R. (2019). A longitudinal neuroimaging dataset on arithmetic processing in school children. *Scientific Data*, 6(190040), 1–14. <https://doi.org/10.1038/sdata.2019.40>
- Suárez-Pellicioni, M., Núñez-Peña, M. I., & Colomé, À. (2016). Math anxiety: A review of its cognitive consequences, psychophysiological correlates, and brain bases. In *Cognitive, Affective and Behavioral Neuroscience* <https://doi.org/10.3758/s13415-015-0370-7>
- Suárez-Pellicioni, M., Prado, J., & Booth, J. (2018). Lack of improvement in multiplication is associated with reverting from verbal retrieval to numerical operations. *NeuroImage*, 183, 859–871. <https://doi.org/10.1016/j.neuroimage.2018.08.074>
- Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different Forces, Same Consequence: Conscientiousness and Competence Beliefs Are Independent Predictors of Academic Effort and Achievement. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/a0017048>
- Updegraff, K. A., Eccles, J. S., Barber, B. L., & O’Brien, K. M. (1996). Course enrollment as self-regulatory behavior: Who takes optional high school math courses? *Learning and Individual Differences*. [https://doi.org/10.1016/S1041-6080\(96\)90016-3](https://doi.org/10.1016/S1041-6080(96)90016-3)
- Weschler, D. (1999). *Weschler Abbreviated Scale of Intelligence*. (The Psyc).
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*. <https://doi.org/10.1006/ceps.1999.1015>
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Riverside Publishing.
- Yates, S. M. (2000). Task involvement and ego orientation in mathematics achievement: A three year follow-up. *Issues In Educational Research*, 10(1), 77-91.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.