Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 04, April : 2022

An FMRI study comparing mathematicians and non-mathematicians found that naive and scientific theories in mathematics and science can interfere with one another.

1.Mrs Soubhagini Mohapatra 2.Dr. Meenakhsi mohanty
Nalanda Institute of Technology, Bhubaneswar
Dept. of Basic Science & Humanities
E-mail ID: soubhagini@thenalanda.com
MeenakshiMohanty@thenalanda.onmicrosoft.com

ARTICLE INFO

Keywords:
Conceptual knowledge
Conceptual interference
Mathematical expertise
Inhibition
fMRI

ABSTRACT

Background: Conceptual interference is one of the most common learning challenges in mathematics. The majority of conceptual interference research has, however, been on science. In this functional magnetic resonance imaging (fMRI) study, we investigated the moderating effects of mathematical expertise as well as the conceptual interference effects in both mathematics and science.
Methods: A speeded reasoning assignment using assertions from mathematics and science was performed by thirty adult mathematicians and 31 non-mathematicians who were matched for gender, age, and intelligence. Statements could be truthful or incorrect in accordance with both naive and scientific conceptions, or they could be incongruent (differed in their truth value).
Findings: In the science and math challenge, both groups made more mistakes and took longer to respond when evaluating incongruent versus congruent statements, but mathematicians were less impacted by naive beliefs. Inhibiting naive hypotheses in mathematics activated the left dorsolateral prefrontal cortex, whereas in science it activated both the dorsolateral and ventrolateral prefrontal cortex bilaterally. The brain level conceptual interference effect was unaffected by mathematical proficiency.
Conclusion: This study shows that naive theories in mathematics still exist among mathematicians, despite the fact that their performance is less impacted by them than that of novices. Additionally, the distinct brain activity throughout the arithmetic and science tasks suggests that the calibre of the concepts being used in the inhibitory control processes determines how much interference is resolved..

---

1. Introduction

The acquisition of academic knowledge is a complex process with a lot of challenges. One of these challenges relates to the presence of hindering misconceptions [1,2]. An increasing body of research has demonstrated that pre-instructional naïve theories can not only impair the acquisition of scientific theories but also interfere with the efficient and accurate processing (e.g., retrieval) of scientifical information (i.e., conceptual interference). This is even the case in individuals who are experts in a specific knowledge domain (for a review see [3]). Despite these findings, the mental and neural mechanisms associated with conceptual interference and its resolution are far from understood, especially in the domain of mathematics. Further, even less is known about how these mechanisms differ in individuals who are experts in a knowledge domain.

In the last decade, it has been proposed that naïve and scientific theories coexist in memory. In their seminal work, Shtulman and Valcarcel [4] used a speeded reasoning task to investigate the performance of well-educated adults across various scientific domains. Partic had to verify 200 scientific statements about different phenomena. These statements belonged to one of two different types: congruent statements were true or false according to both scientifically and naïve theories (e.g., “The moon revolves around the earth”), incongruent statements, in contrast, differed in their truth value between the theories (e.g., “The sun revolves around the earth”: naively true but scientifically false). The authors observed that incongruent statements were verified more slowly and less accurately than congruent statements. This behavioral pattern indicates that even after the acquisition of scientifi- cally correct knowledge naïve theories still exist in memory and that a cognitive conflict occurs when incongruent statements are verified, resulting in poorer performance.

Since its discovery, the conceptual interference effect has been replicated in numerous studies (e.g., [3–7]) and has also been found in individuals with high scientific knowledge or expertise (e.g., [8–11]). For instance, Shtulman and Harrington [5] asked science professors with three to four decades of scientific expertise (n = 10), humanities professors (n = 11), and similarly aged adults from the local community (n = 27) to perform the speeded reasoning task described above. Science professors performed significantly better in the various domains, demonstrating overall higher accuracy and shorter response times than the two other groups. Nevertheless, the science professors were slower and made more errors in the incongruent condition compared to the congruent condition. More importantly, in comparison to the other two groups, they showed a smaller difference in accuracy between incongruent and congruent statements (9%) compared to humanities professors (13%) and non-professors (20%). Although this pattern was only found for the accuracy data, the results suggest that although individuals with scientific expertise continue to be influenced by naïve theories, they are affected to a lesser extent than individuals with lower expertise.

Compared to the considerable body of evidence outlined above, research on conceptual interference in the domain of mathematics is very scarce. This may appear surprising since mathematics is an important part of everyday life and serves as a foundation for various scientific and technological domains. While Shtulman and colleagues [4, 5] included some mathematical misconceptions (i.e., fractions), a more comprehensive investigation of the influence of incorrect naïve mathematical theories was recently done by Stricker et al. [6]. The authors developed a mathematical adaptation of the speeded reasoning task from Shtulman and Valcarcel [4], consisting of 49 congruent and incongruent problem sets from various mathematical sub-domains: including fractions, algebra, units and geometry, probability, and basic numerical concepts. The authors presented these problems to 62 university students. As expected, participants demonstrated slower reaction times and more errors for incongruent compared to congruent statements across all five mathematical sub-domains. To further examine the relation between mathematical competence and interference, the authors correlated several measures of mathematical achievement (mathematical competence, arithmetic fluency, and math grade) with interference indices in accuracy and reaction times (i.e., the difference between congruent and incongruent trials). This correlational analysis revealed several significant associations between the accuracy interference index and mathematical competence (−0.69), math grade (0.40), and arithmetic fluency (−0.29), reflecting that individuals with higher mathematical skills exhibit a smaller interference effect. Thus, this work suggests that naïve mathematical theories continue to influence the performance during mathematical judgments and that the individual level of conceptual interference is negatively associated with mathematical achievement.

While the study by Stricker and colleagues [6] indicates that mathematical competence is reducing conceptual interference in mathematics, it does not tell us whether this also holds true for mathematical expertise. This question has been investigated in two other behavioral studies [11,12]. Obersteiner and colleagues [11] presented 44 adult mathematicians (holding at least a Master degree) with a fraction comparison task related to the natural number bias (i.e., focusing on the natural numbers in evaluating the size of fractions). The fraction pairs were either congruent (the natural number bias results in a correct response) or incongruent (the natural number bias would result in an incorrect answer). The results confirmed their expectations that even experts show longer response times for incongruent compared to congruent trials. Inconsistent with their expectations, the authors did not find differences in accuracy (probably due to a ceiling effect of 98% accuracy). However, the authors did not compare experts to novices in the field. In another behavioral work, Lubin and colleagues [12] compared 25 students of mathematics with 37 intelligence-matched students of humanities. The authors used arithmetic word problems in which the relational term (e.g., “more than” or “less than”) could be either congruent (addition) or incongruent (subtraction) with the required arithmetic operation. The results demonstrated that incongruent word problems were associated with longer response times and lower accuracies in both groups. The students of mathematics, however, were less affected by naïve theories, as reflected in small interference effects in both accuracy and response time. Thus, the overall pattern indicates that naïve mathematical theories can interfere with the process of finding a (scientifically) correct answer even in individuals with mathematical expertise.

A prominent hypothesis to explain the impact of naïve theories is that they need to be actively inhibited to resolve the occurring cognitive interference so that the scientific theory can determine the response (for a recent review, cf. [3]). The involvement of inhibition mechanisms during conflict resolution has been observed in several neuroimaging studies. More specifically, they have demonstrated that incongruent conditions/statements engage brain regions that have been associated with inhibitory control processes, i.e., the dorsolateral prefrontal cortex (DLPFC, Brodmann Areal (BA) 8,9,46) and the ventrolateral prefrontal cortex (VLPFC, BA 44,45,47), to a larger extent than congruent conditions/statements [7,13]. Some of these studies also found a stronger activation of the anterior cingulate cortex (ACC, BA 24,32), a brain region that is associated with error detection and conflict monitoring [9, 14,15]. These findings corroborate the notion that inhibition might play an important role in suppressing irrelevant or false naïve information. To date, only one neuroimaging study has investigated the brain mechanism in the domain of mathematics. Stavy and Babai [1] used functional magnetic resonance imaging (fMRI) to measure the brain activation of 14 university students in relation to a specific misconception in mathematics: the belief that a shape with a larger area also has a larger perimeter and that shapes with the same perimeter have the same area. In the scanner, students were asked to compare the perimeter of two geometrical shapes, using a congruent and incongruent condition. Besides replicating the behavioral results of previous work (i.e., longer response times and lower accuracy for incongruent items), the authors showed that incongruent trials activated bilateral areas of the prefrontal cortex, including parts of the VLPFC (BA 10/11 and BA 11/47), more strongly than congruent trials. Further, the individual interference effect in response time was negatively related to the brain activation of the left prefrontal area (VLPFC, BA 45). Overall, these findings indicate that conceptual interference in mathematics might engage similar brain regions as other domains and that individuals who demonstrate lower interference activate this brain area more strongly. Thus, it appears plausible to conclude that experts are more proficient in inhibiting naïve domain-specific information and experience a reduced cognitive conflict. Indeed, some evidence from neuroimaging studies in the domain of science is pointing in this direction [9,10].

For instance, Allaire-Duquette and colleagues [9] investigated the brain activation of 25 physics experts with a PhD who were presented with speeded-reasoning tasks in two different domains (physics and biology). Besides a poorer performance during incongruent statements in both domains, the fMRI analysis revealed higher brain activation for incongruent compared to congruent statements in the left inferior frontal gyrus (part of the VLPFC), the bilateral superior frontal gyrus(part of the DLPFC), and the bilateral ACC. These findings converge with the brain regions reported above, indicating that even in experts cognitive inhibition mechanisms might play an important role to overcome naïve theories (for similar findings in the domain of chemistry, see [10]). Nevertheless, the present findings are also limited because a direct comparison to non-experts (i.e., novices) was not performed. Hence, it is not known whether the observed brain activations within these domains are specific to experts or not. Some brain imaging studies have directly compared the brain activation of experts and novices but did not directly contrast incongruent versus congruent conditions [14, 16], which isolates possible inhibition processes. The interpretation of the results is further complicated as the behavioral responses between experts and novices differed substantially (experts answered mostly correctly, novices incorrectly, [14]). Thus, it remains unclear whether the observed brain activation differences can be attributed to inhibition mechanisms per se or to error detection and error monitoring (or other cognitive processes that have been associated with these brain regions).

To summarize, research on conceptual knowledge representation in mathematics is sparse. Further, there are only a few studies that have examined the role of mathematical expertise, and only one that

Table 1

Descriptive statistics and frequentist statistics (t-test for independent samples) for mathematicians and non-mathematicians.

Variable	Mathematicians *M* (*SD*)	Non-mathematicians *M* (*SD*)	Independent samples *t*-test	Effect size Cohen's d
Age (years)	23.23 (3.66)	24.13 (4.29)	*t*(59) = —0.88, *p* = .384	*d* = —0.23
Experience (years)	4.07 (3.40)	4.29 (3.02)	*t*(59) = —0.27, *p* = .787	*d* = —0.07
General intelligence (mean z-score)	0.06 (0.44)	—0.06 (0.47)	*t*(58) = 1.01, *p* = .317	*d* = 0.26
Numerical Intelligence (mean z-score)	**0.17 (0.63)**	**¡ 0.15 (0.60)**	***t*(58) ¼ 2.01,** ***p* = .049**	***d* ¼ 0.52**
Verbal intelligence (mean z-score)	—0.13 (0.53)	0.12 (0.67)	*t*(58) = —1.60, *p* = .115	*d* = —0.41
Figural Intelligence (mean z-score)	0.15 (0.63)	—0.14 (0.55)	*t*(58) = 1.88, *p* = .065	*d* = 0.49
Math achievement (raw score)[a]	**28.52 (3.11)**	**20.90 (5.21)**	***t*(49.56) ¼ 7.61,** ***p* < .001**	***d* ¼ 1.76**

Variables, where the t-test showed significant group differences, are bolded

[a] Levene's test is significant, therefore a Welch test was calculated.

investigated the neural correlates associated with a specific mathematical misconception. Therefore, the first aim of this study is to provide a comprehensive picture of conceptual interference effects in mathematics (across different topics) and their dependence on expertise. To this end, we compare the performance of experts and novices in mathematics in a speeded-reasoning task drawing on broad mathematical knowledge. A speeded-reasoning task in the science domain will be administered as a control task. Additionally, while some results suggest that better inhibitory control mechanisms may provide an advantage in overcoming misconceptions in mathematics, there is only one neuroimaging study on one specific mathematical misconception with a quite small sample size (n=14) supporting this assumption [1]. Thus, the second purpose of this study is to investigate the brain activation of mathematicians and non-mathematicians during the evaluation of mathematical and science statements. Finally, another limitation of the majority of the existing behavioral studies comparing experts and novices and in all respective neuroscientific studies is that it is unclear if observed group differences were solely due to differences in their domain-specific knowledge or confounded with differences in domain-general abilities. As intelligence is a serious confound, we matched mathematicians and non-mathematicians in terms of generalintelligence.

Based on the literature reviewed above, the following hypotheses (H) are tested. In the domain of mathematics, we expect that incongruent statements are solved less accurately and more slowly compared to congruent ones (H1). This should be the case for both mathematicians and non-mathematicians. Mathematicians should solve the mathematical statements more accurately and faster compared to non-mathematicians (H2). Further, the interference effect (between incongruent and congruent statements) in accuracy should be smaller in mathematicians compared to non-mathematicians (H3). Since the results regarding response time are inconsistent, we do not make a specific assumption for them. In the domain of science, we also expect lower accuracy and longer response times in incongruent compared to congruent statements (H4). Since the two groups should not differ in science expertise, we expect group differences neither in overall performance (H5) nor in the interference effect (H6). The fMRI data will provide additional evidence on inhibitory control mechanisms and specific differences in expertise. We hypothesize that the DLPFC, VLPFC, and ACC are activated more strongly while evaluating incongruent than congruent statements in both domains (H7). Further, we assume that mathematicians show stronger activation in the DLPFC, VLPFC, and ACC compared to non-mathematicians in the contrast incongruent > congruent mathematical statements, indicating greater recruitment of inhibitory control mechanisms (H8). For the science statements, we do not expect a significant group difference in the neural interference effect (H9).

2. Method

Participants

For this study, 61 adults were recruited from the University of Graz.30 of them (24 men, 6 women) were mathematicians, the other 31 (23men, 8 women) were non-mathematicians. Mathematicians = (n 30)were defined as individuals who study or have studied mathematics atthe university. Non-mathematicians (n = 31) were defined as in-dividuals who study or have studied a subject with no to minimalexplicit mathematical content. Non-mathematicians were recruitedfrom the following subjects: Teaching (different subjects; $n = 9$), Med-icine ($n = 7$), Translation ($n = 4$), Musicology ($n = 2$), Law ($n = 1$),History ($n = 1$), German Philology ($n = 1$), Archeology (v 1), DentalMedicine ($n = 1$), Geography ($n = 1$), Art History ($n = 1$), Sport Studies($n = 1$), and Sustainable Development ($n = 1$). All individuals wereeither currently completing their Bachelor's degree (Mathematicians: n24; Non- = = mathematicians: n 19), their Masters' degree (Mathema- = = ticians: n 4; Non-mathematicians: n 9), or their PhD = = (Mathemati-cians: n 2; Non-mathematicians: n 3). Both groups were matchedaccording to sex, age, professional experience (years spent studying andworking in their field of expertise), and general intelligence (see Table 1).

To ensure that intelligence-matched mathematicians and non-mathematicians differed only in their amount of mathematical expertise, we compared their performance in specific intelligence domains (numerical, verbal, and figural) and their performance in a mathemat- ical achievement test, measuring higher-level math competencies (more information on both measurements can be found under 2.2.3. and2.2.4.). As ensured by the matching procedure, mathematicians differed substantially in mathematical achievement, and, to a small extent, in numerical intelligence from non-mathematicians. However, the two groups neither differed in demographic characteristics nor in general intelligence, nor in the other two intelligence domains (see Table 1).

All participants had German as their native language, were right-handed, and had normal or corrected to normal vision. The participants neither reported a history of psychiatric, neurological, or learning disorders, nor a current use of psychoactive medication. All participants gave informed consent and were compensated with a total of 30 € for

1.5 h of participation. The experimental procedure of the study was approved by the local ethics committee.

Material

Conceptual interference in mathematics

To measure interference and cognitive mechanisms to overcome interference, participants had to decide whether a statement is

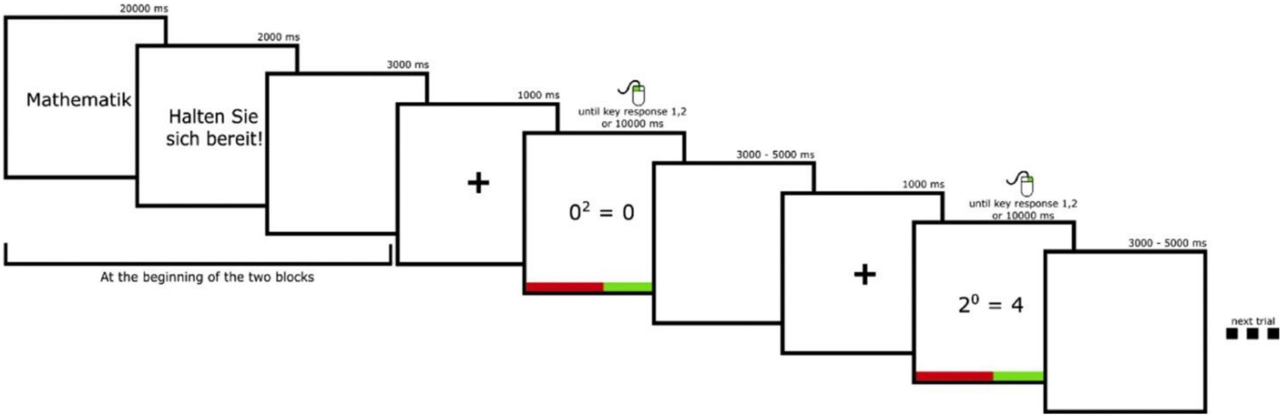


Fig. 1. Example trial and presentation times of the interference of naïve and mathematical theories task.

mathematically correct or incorrect. The statements were developed by experts in mathematics education and include frequent misconceptions in mathematics (for a full description of the items see [6]). From the original 196 statements of this work [6], we chose 100 statements that elicited the largest interference in response time and accuracy using unpublished data from 153 individuals. Topics of those statements were fractions, algebra, units and geometry, and basic concepts (a full list of all statements can be found in Table A1 in the Appendix). Those 100 items assess 25 different mathematical concepts with 4 items each. One item is mathematically and naively true (e.g., 1/4 + 2/4 = 3/4), one item is mathematically and naively false (e.g., 1/4 + 1/4 = 1/4), one item is mathematically false and naively true (e.g., 1/3 + 1/4 = 2/7), and one item is mathematically true and naively false (e.g., 1/10 + 1/10 = 1/5). While the first two items are congruent across naïve and mathematical theories, the latter two are incongruent and should induce interference and cognitive mechanisms to overcome interference. For each statement accuracy and response times were collected in addition to brain activation.

Conceptual interference in science

As a control task, participants had to work on a conceptual similar task, only with scientific statements. Participants had to decide for 100 statements whether the statement is scientifically correct or incorrect. The statements were chosen from the 200 statements used in the seminal study by Shtulman and Valcarcel [4]. Again, we chose those statements that elicited the largest interference in response time and accuracy using unpublished data from 153 individuals. Topics of those statements were astronomy, evolution, genetics, germs, matter, mechanics, physiology, thermodynamics, and waves (a full list of all statements can be found in Table A2 in the Appendix). Those 100 items assess 25 different scientific concepts with 4 items each. One item is scientifically and naively true (e.g., fish are alive), one item is scientifically and naively false (e.g., stones are alive), one item is scientifically false and naively true (e.g., the sun is alive), and one item is scientifically true and naively false (e.g., corals are alive). While the first two items are congruent across naïve and scientific theories, the latter two are incongruent and should induce interference. For each statement accuracy and response time data were collected in addition to brain activation.

Berlin intelligence structure test, short version (BIS-T)

To assess intelligence we used the short version of the Berlin Intelligence Structure Test [17]. The short version is a structured paper pencil test that consists of 15 tasks, each task is a different combination of one of the three domains of intelligence (numerical, verbal, and figural) as well as of one of the four operational abilities (processing speed, memory, reasoning, and creativity). The internal consistencies of the separate tasks are considered appropriate (Cronbach’s α 0.75 - 0.89). The processing time for all 15 tasks was approximately 45 min. We used z-standardized raw scores for our analyses. The respective z-scores were averaged for each of the three subscales (numerical, verbal, figural) and all z-scores were averaged for a general intelligence score.

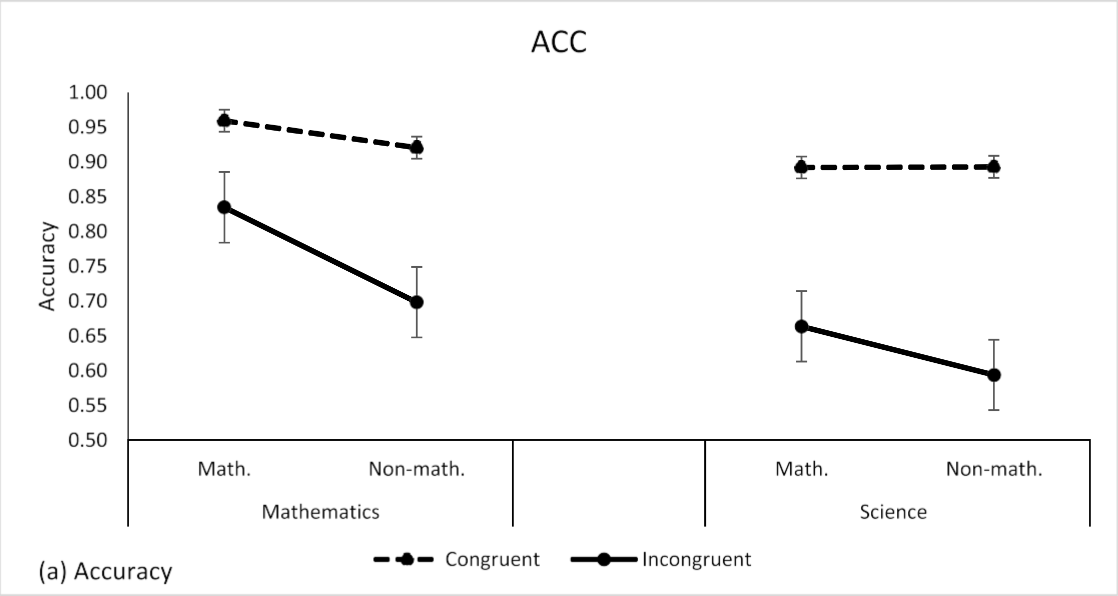
Mathematics test for selection of personnel (M-PA)

To assess mathematical achievement, we used the mathematics testfor selection of personnel [18]. This is a paper-pencil test originallyconstructed to assess mathematical abilities for job applications andmeasures performance in higher-order mathematics including fractions,conversion of units, exponentiation, division with decimals, algebra,geometry, roots, and logarithm. We used the official short version,consisting of 31 mathematical problems. The short version has a goodinternal consistency (Cronbach's α = = 0.89) and correlates very high (r 0.93) with the long version [18]. The processing time is limited to 15min and we used individuals' raw scores of all correct answers (rangingfrom 0 to 31) for our analyses.

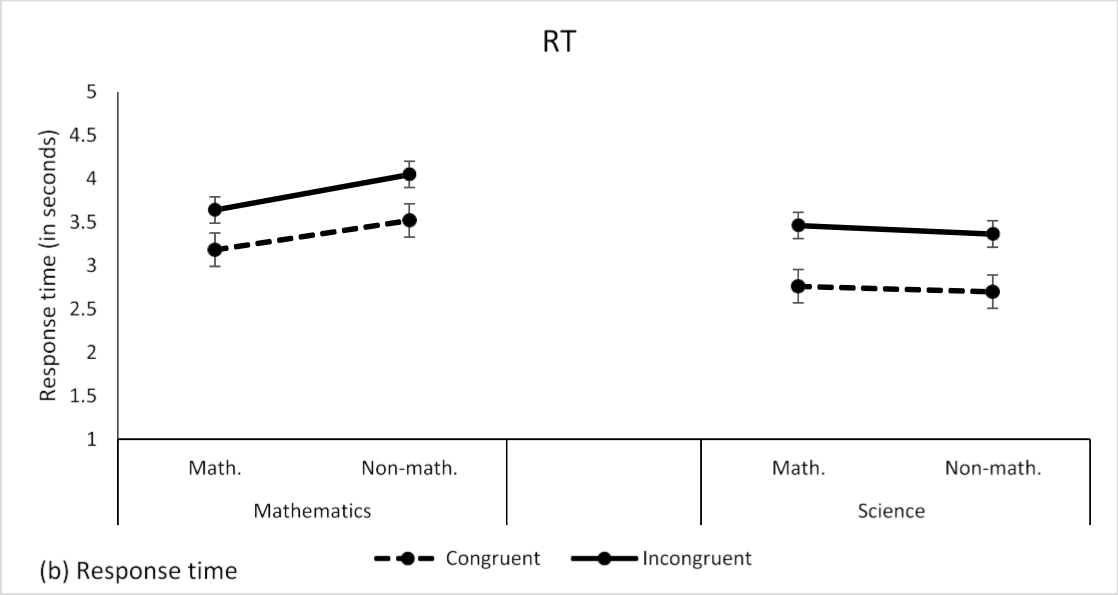
Procedure

More than half of the participants (60%) took part in a previous study [19], in which their intelligence scores were obtained. The intelligence scores of the remaining participants (40%) were collected in an addi- tional, separate session. One mathematician did not come back for the intelligence measurements leading to missing data. The intelligence measurements were obtained either before (N = 48; M = 241, SD = 153, $Range$ = 5–750 days) or after (N = 12; M = 159, SD = 100, $Range$ = 31–302 days) the fMRI session.

The procedure of the fMRI session was as follows. After the participants were welcomed and had to fill out a COVID-19 safety checklist, they all received information regarding the task and the fMRI scan and had to give informed consent. Before being placed into the scanner, there was a practice run to familiarize the participants with the task and the response box. After participants were placed in the scanner, a T1 structural scan was made. The experimental tasks (speeded reasoning tasks) were programmed with Psychopy3 [20] and were presented to the participants while lying in the scanner using a monitor and an overhead mirror. Participants started either with the mathematical or with the scientific statements. Before each run participants were informed on the computer screen (20 s) whether mathematical or scientific statements will be presented (see also Fig. 1). Within each domain, all 100 itemswere presented in randomized order. Each item started with a black fixation cross on white background for 1 s. The statements were pre- sented for a maximum of 10 s or until an answer was given. Half of the participants had to press the left button of the response box with their index finger for correct and the right button of the response box with their middle finger for incorrect. The other half of the participants had to press the left button of the response box with their index finger for incorrect and the right button of the response box with their middle



(a) Accuracy



(b) Response time

Fig. 2. Means and standard errors of (a) accuracy; (b) response time in seconds for mathematical and science statements separated for mathematicians (Math.) and non-mathematicians (Non-math.).

finger for correct. Below each statement a red and green box indicated which button to use to judge the statement as correct and which to judge as incorrect. Each statement was followed by a jittered interstimulus interval with a mean duration of 4 s (range between 3 and 5 s).

Due to the self-paced paradigm, the duration was different for each participant, however the majority of all participants needed between 28 and 33 min (*M*=30.28, *SD* =2.44) to complete this task. There was no significant difference between mathematicians (*M* =30.22, *SD* = 2.20) and non-mathematicians (*M* =30.34, *SD* =2.68, *p* =.854) in task duration. After finishing the task, a diffusion tensor imaging (DTI) scan and a functional resting state (rs-fMRI) scan were made (15 min). Overall, the participants spent between 45 and 60 min in the scanner, the complete study took between 60 and 75 min. After leaving the scanner, participants were reimbursed and received images of their brain.

MRI protocol

Structural and fMRI data were collected with a 3-T Vida (Siemens)and a 64-channel head coil at the MRI Lab of Graz. Functional imageswere obtained using interleaved gradient echo-planar imaging (EPI)functional images (TR = 1800 ms, TE = 30 ms, FoV = 220 mm, flip angle83°, slices = 72, 2.0 × 2.0 × 2.0 mm isotropic voxel resolution). The structural images were obtained using a high-resolution single shot T1-weighted anatomical image with a generalized autocalibrating partiallyparallel acquisitions (Grappa) sequence (TR = 1600 ms, TE = 2.38 ms,FoV = 224 mm, flip angle 9°, 1.0 × 1.0 × 1.0 mm isotropic voxel resolu-tion). The scan time of the structural images was 4 min. Furthermore, DTI and functional resting state (rs-fMRI) data were also acquired butnot analyzed in the present study.

Analyses

Behavioral analyses

Behavioral data was analyzed using IBM SPSS Statistics [21]. First, 2×2 ANOVAs with the between-groups variable group (mathematicians vs. non-mathematicians) and the within-group variable congruency (congruent vs. incongruent) were calculated. The dependent variables were accuracy as well as response time in the mathematical task and in the science task, resulting in four ANOVAs (H1, H2, H4, H5). To investigate whether mathematicians or non-mathematicians are more affected by interference induced by incongruent statements, we also operationalize the amount of interference (H3, H6). We calculated four interference indices after Stricker et al. [6]: one index for accuracy (1) and one for response time (2) separately for the mathematical and the science tasks.

$$\text{interference}_{ACC} = \text{congruent statements}_{ACC} - \text{incongruent statements}_{ACC} \quad (1)$$

$$\text{interference}_{RT} = \text{incongruent statements}_{RT} - \text{congruent statements}_{RT}(\text{correct responses only}) \quad (2)$$

Four independent t-tests were calculated to test whether the above interference indices for accuracy (1) and response time (2) differ between mathematicians vs. non-mathematicians in the mathematical task and the science task. The behavioral data that support the findings of this study are openly available in OSF athttps://osf.io/vhcte/?view_only=3d508a8c0998454cbcc5d3769c500c16.https://osf.io/vhcte/?view_only=82648f3dafa34d7c9d13522e398044ff

Image processing and functional analyses

The preprocessing of the functional imaging data was done using fMRIPrep [22]. FMRIPrep is a robust and reproducible data processing pipeline designed to provide state-of-the-art preprocessing. Within this pipeline the functional imaging data were motion corrected, slice-time corrected, co-registered with the structural T1 data, segmented, and then normalized to the MNI template. We used a Gaussian filter of 9 mm full-width-at-half-maximum (FWHM) to further smooth the preprocessed imaging data with the software package FitLins [23].

In the present work, we implemented a two-level approach to analyze the imaging data. First, we calculated a general linear model (GLM) for each participant (first-level analyses) with FitLins, including additional covariates for translation, rotation, framewise displacement, and anatomical correlates. We then estimated a second-level random effects whole-brain model using the Matlab Toolbox GLM_Flex_Fast4 [24]. To statistically test interference related brain activations associated with mathematical and scientific statements, we calculated the contrasts between "incongruent > congruent" statements for each of the domains (mathematics and science; H7). We then contrasted "incongruent > congruent" statements of each domain (mathematics and science) in a first step within the groups of mathematicians and non-mathematicians. In a second step, we performed a contrast analysis between the groups to evaluate expert-related interference effects (H8, H9; for more information see Appendix 6.3.). All reported t-statistics were family-wise error (FWE) corrected with a threshold at $p_{FWE} < 0.001$ at voxel level. The threshold correction and the figures were done using the Nilearn toolbox in Python. For the automatic labeling of the brain regions and the Brodmann Areas we used the label4MRI R package [25] which uses the AAL (automatic anatomical labeling) atlas [26].

Table 2
Independent samples *t*-test for the interference indices comparing mathematicians (Math.) and non-mathematicians (Non-math.).

Variable	Math. *M* (*SD*)	Non-math. *M* (*SD*)	Independent samples *t*-test	Effect size Cohen's d
Mathematics: interference$_{ACC}$ [a]	**0.13 (0.06)**	**0.22 (0.10)**	***t*(48.88) ¼ ¡ 4.88, *p* < .001**	***d* ¼ ¡ 1.24**
Mathematics: interference$_{RT}$	0.46 (0.23)	0.53 (0.30)	*t*(59) = — 1.05, *p* = .297	*d* = — 0.27
Science: interference$_{ACC}$	**0.23 (0.10)**	**0.30 (0.10)**	***t*(59) ¼ ¡ 2.82, *p* = .006**	***d* ¼ ¡ 0.72**
Science: interference$_{RT}$	0.70 (0.32)	0.61 (0.36)	*t*(59) = 0.40, *p* = .668	*d* = 0.10

Variables, where the *t*-test showed significant group differences, are bolded.
[a] Levene's test is significant, therefore a Welch test was calculated.

3. Results

Behavioral

Conceptual interference in mathematics

The analysis on accuracy data showed significant main effects of congruency ($F(1,59) = 294.47$, $p < .001$, $\eta_p^2 = .83$) and group ($F(1,59) = 32.78$, $p < .001$, $\eta_p^2 = .36$). In line with previous evidence and as expected (H1), congruent statements were solved more accurately than incongruent statements, and mathematicians solved the mathematical statements with a higher accuracy than non-mathematicians (H2). Additionally, there was a significant interaction of congruency and group ($F(1,59) = 23.41$, $p < .001$, $\eta_p^2 = .28$; see Fig. 2a). Both mathematicians ($p < .001$, $d = 1.74$) and non-mathematicians ($p < .001$, $d = 3.11$) showed a congruency effect in that they solved congruent statements more accurately compared to incongruent statements. However, the difference between congruent and incongruent statements, as calculated by the interference index (see Table 2), is smaller in mathematicians (13% lower accuracy for incongruent statements) compared to non-mathematicians (22% lower accuracy for incongruent statements). This confirms our third hypothesis (H3) and indicates that even though domain experts are still affected by naïve concepts, they are less affected by the conflict between naïve and mathematical theories than novices. Additionally, the difference in accuracy between mathematicians and non-mathematicians is larger in the incongruent condition ($p < .001$, $d = 1.91$) compared to the congruent condition ($p = .001$, $d = 0.54$). This indicates that mathematicians have an advantage especially when dealing with interfering naïve and mathematical theories.

The analysis on response times revealed significant main effects of congruency ($F(1,59) = 212.68$, $p < .001$, $\eta_p^2 = .78$) and group ($F(1,59) = 4.16$, $p = .046$, $\eta_p^2 = .07$). As expected, congruent statements were solved faster than incongruent statements (H1), and the mathematicians solved the mathematical statements faster than the non-mathematicians (H2; see Fig. 2b). In line with our hypothesis (H3), there was no significant interaction of congruency and group ($F(1,59) = 1.11$, $p = .297$, $\eta_p^2 = .02$), as well as no significant group difference in the interference index (see Table 2). This indicates that the interference between naïve and mathematical theories affects response speed of experts and novices to the same extent.

Conceptual interference in science

In the domain of science, the analysis on accuracy data yielded sig-

nificant main effects of congruency ($F(1,59) = 444.56$, $p < .001$, η_p^2
=
.88) and group on accuracy ($F(1,59)$ 4.25, p .044, η_p^2 .07). Again, = = =
as expected, incongruent statements were solved less accurately than congruent statements (H4). Contrary to our hypothesis (H5) mathema- ticians solved the science statements with higher accuracy than non- mathematicians. Additionally, we found a significant interaction effecton accuracy ($F(1,59) = 7.97$, $p = .006$, $\eta_p^2 = .12$). Post-hoc pairwise

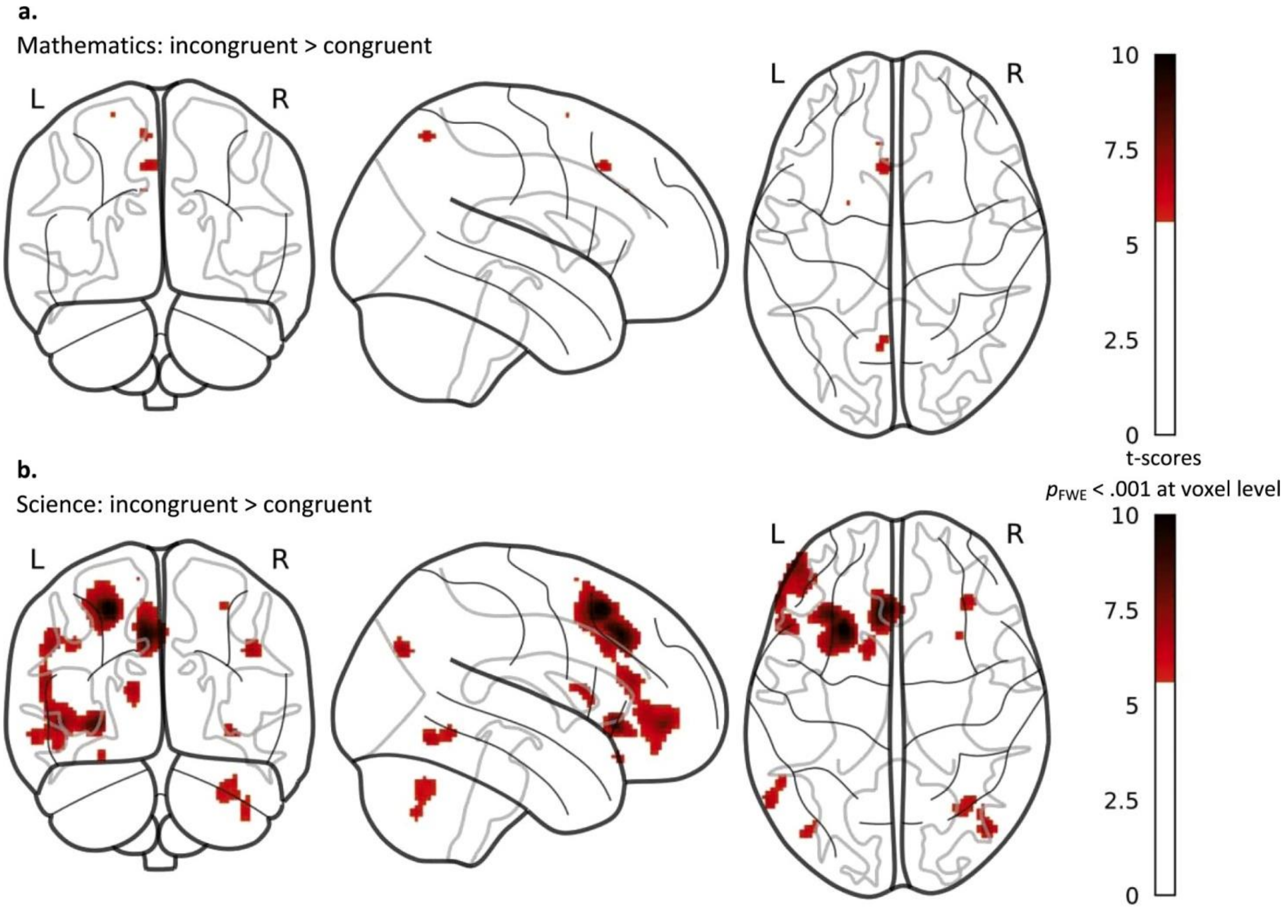


Fig. 3. **a.** Mathematics: incongruent > congruent. **b.** Science: incongruent > congruent.

comparison with Bonferroni correction for multiple comparison showed that both mathematicians ($p < .001$, $d = 2.81$) and non-mathematicians ($p < .001$, d=3.67) displayed a congruency effect and solved congruent statements more accurately compared to incongruent statements (see Fig. 2a). In contrast to our hypothesis (H6), the accuracy difference between congruent and incongruent statements, as calculated by the interference index, is smaller in mathematicians (Table 2). Mathematicians solved incongruent statements with a 23% lower accuracy than

Table 3
One-sided related samples t-test statistics for the contrast incongruent > congruent for each of the domains.

Brain Region	AAL Label	Brodmann Areal (BA)	MNI X	Y	Z	Cluster size k
(A) Mathematics: Incongruent > Congruent						
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	-5	20	44	200
Superior parietal gyrus	Precuneus_L	Left-BA7	-7	-63	58	104
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	-9	32	32	16
(B) Science: Incongruent > Congruent						
Lateral DLPFC	Frontal_Mid_L	Left-BA8	-25	20	52	4176
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	-7	30	40	3040
Insula	Insula_L	Left-BA13	-31	26	-3	1160
Anterior prefrontal cortex	Frontal_Mid_Orb_L	Left-BA10	-47	50	-3	3424
Lateral DLPFC	Frontal_Inf_Tri_L	Left-BA46	-55	36	18	
VLPFC	Frontal_Inf_Orb_L	Left-BA47	-49	46	-15	
Angular gyrus	Occipital_Mid_R	Right-BA39	44	-75	32	376
Lateral DLPFC	Frontal_Inf_Oper_L	Left-BA9	-53	24	34	744
Caudate	Caudate_L	Left-BA48	-13	12	14	352
VLPFC	Frontal_Inf_Orb_R	Right-BA47	34	34	-7	184
Inferior temporal gyrus	Temporal_Inf_L	Left-BA37	-61	-63	-11	448

Inferior temporal gyrus	Temporal_Inf_L	Left-BA37	-53	-53	-9	
VLPFC	Frontal_Inf_Orb_L	Left-BA47	-29	28	-19	80
Angular gyrus	Occipital_Mid_L	Left-BA39	-43	-77	34	216
Cerebellum crus I	Cerebelum_Crus1_R	Right-Cerebellum	30	-67	-37	536
Cerebellum crus I	Cerebelum_Crus1_R	Right-Cerebellum	36	-63	-31	
Cerebellum crus II	Cerebelum_Crus2_R	Right-Cerebellum	42	-69	-45	160
Lateral DLPFC	Frontal_Mid_R	Right-BA8	30	16	54	56

All comparisons are reported at FWE corrected threshold at $p_{\mathrm{FWE}} < 0.001$ at voxel level; Clusters with a cluster size smaller than 10 voxels are not reported.

Table 4
One-sided related samples t-test for the contrast incongruent > congruent for the domain science, separately for mathematicians and non-mathematicians.

Brain Region	AAL Label	Brodmann Areal (BA)	MNI			Cluster size k
			X	Y	Z	
(A) Mathematicians: Incongruent > Congruent in Science						
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	—7	30	40	944
Lateral DLPFC	Frontal_Mid_L	Left-BA8	—27	20	52	1208
Caudate	Caudate_L	Left-BA48	—13	10	14	320
Lateral DLPFC	Frontal_Mid_L	Left-BA9	—51	24	36	440
Lateral DLPFC	Frontal_Inf_Tri_L	Left-BA46	—55	34	18	304
Anterior prefrontal cortex	Frontal_Mid_Orb_L	Left-BA10	—49	48	—3	584
Cerebellum Crus I	Cerebelum_Crus1_R	Right-Cerebellum	36	—63	—29	64
Inferior temporal gyrus	Temporal_Inf_L	Left-BA37	—61	—63	—11	56
Cerebellum Crus I	Cerebelum_Crus1_L	Left-Cerebellum	—29	—69	—31	64
VLPFC	Frontal_Inf_Orb_L	Left-BA47	—35	26	—5	32
(B) Non-mathematicians: Incongruent > Congruent in Science						
Insula	Insula_L	Left-BA13	—29	26	—3	472
Lateral DLPFC	Frontal_Mid_L	Left-BA8	—25	20	52	792
Supplementary motor area	Frontal_Mid_L	Left-BA6	—29	10	44	
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	—9	24	44	376
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	—7	30	38	

All comparisons are reported at FWE corrected threshold at $p_{FWE} < 0.001$ at voxel level; Clusters with a cluster size smaller than 10 voxels are not reported.

congruent statements, non-mathematicians showed a larger difference of 30%, indicating that mathematicians are less affected by naïve theories in science compared to the non-mathematicians. Further, Fig. 2a shows that while mathematicians solved the incongruent statements more accurately than the non-mathematicians ($p = .011$, $d = 0.86$), the congruent statements were solved equally accurately by mathematicians and non-mathematicians ($p = .943$, $d = —0.01$). These post-hoc results show that the small difference in overall accuracy is only due to the difference in the incongruent condition. These results indicate that mathematicians seem to have an advantage compared to non-mathematicians in dealing with interfering naïve theories also in the domain of science.

The analysis on response times showed a significant main effect of congruency ($F(1,59) = 245.35$, $p < .001$, $\eta_p^2 = .81$) but no significant main effect of group ($F(1,59) = 0.22$, $p = .645$, $\eta_p^2 < .01$). As hypothesized, congruent statements were solved faster than incongruent statements (H4), and mathematicians solved the scientific statements equally fast as the non-mathematicians (H5; see Fig. 2b). In addition,

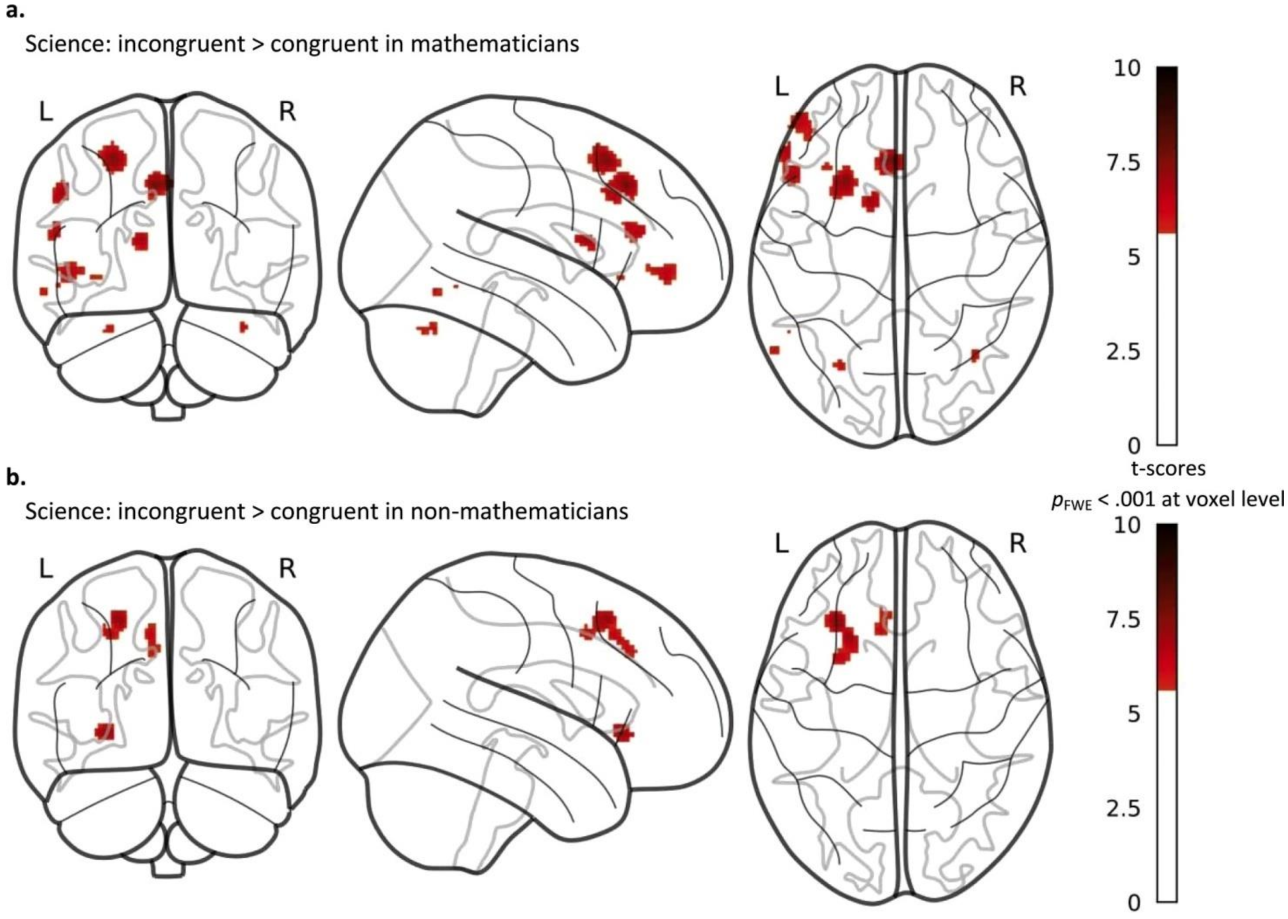


Fig. 4. **a.** Science: incongruent > congruent in mathematicians. **b.** Science: incongruent > congruent in non-mathematicians.

there was no significant interaction of congruency and group ($F(1,59) = 0.16$, $p = .688$, $\eta_p^2 < 0.01$), indicating that with regards to response time both mathematicians and non-mathematicians were affected to the same extent by interfering naïve and mathematical theories (H6). This is also confirmed by a non-significant between group difference in the interference index (Table 2).

fMRI

Overall interference effects in the mathematical task and the sciencetask

We first examined which regions showed an interference related modulation (incongruent > congruent) in association with the mathematical statements. The results revealed significantly greater activation for incongruent compared to congruent statements in the left medial DLPFC and the left medial section of the superior parietal gyrus, more specifically in the precuneus (see Fig. 3a and Table 3a).

Results for the science task revealed several significant brain regions that showed a larger engagement during incongruent compared to congruent statements (see Fig. 3b and Table 3b). These regions included the bilateral DLPFC, left medial DLPFC, and bilateral VLPFC. Contrary to our expectations, we found no significant activation differences in the ACC, neither in the mathematical task nor in the science task. Thus, our hypothesis (H7) was only supported with respect to the DLPFC in the mathematical task and the DLPFC and VLPFC in the science task.

Interference effects of mathematicians and non-mathematicians inthe mathematical task

As a next step, we analyzed interference related brain activation (incongruent > congruent) for mathematical statements in mathematicians and non-mathematicians separately. Contrary to our hypothesis (H8), this group-wise analysis did not reveal any significant activation differences, neither in mathematicians nor in non-mathematicians.
Further, we also did not find significant activation differences in the contrast incongruent > congruent between the groups (see Table A3 in the appendix). An exploratory conjunction analysis showed also no results (see Table A3 in the appendix).

Interference effects of mathematicians and non-mathematicians inthe science task

In the domain of science, we observed several interference related activation clusters (incongruent > congruent) for the entire sample. Since mathematicians and non-mathematicians were not assumed to differ in their science expertise, we also did not expect group differences in this contrast (H9). However, in contrast to the mathematical task, we found several significant activation clusters in both groups (see Table 4 and Fig. 4). Mathematicians showed a greater interference related brain activation (incongruent > congruent) in the left lateral and left medial DLPFC, the left VLPFC, the left caudate, the left anterior prefrontal cortex, the cerebellum bilaterally, and the left inferior temporal gyrus. In comparison, non-mathematicians showed a greater interference related activation in the left lateral and left medial DLPFC, the left insula, and the left supplementary motor area. The contrast analysis showed no significant brain areas in which the incongruent > congruent contrast differed between groups (see Table A4 in the appendix). An exploratory conjunction analysis revealed that both mathematicians and non-mathematicians demonstrated significant activation in the lateral and medial DLPFC (see Fig. A1 and Table A4 in the appendix).

4. Discussion

The present work is the first study to investigate conceptual interference in experts and novices within the domain of mathematics from a neurocognitive perspective. We compared 30 mathematicians with 31 intelligence matched non-mathematicians in two speeded reasoning tasks from the domains of mathematics and science while measuring brain activation using fMRI. Replicating previous behavioral findings, both groups showed higher error rates and longer response times for incongruent statements compared to congruent statements. This was observable in the domain of mathematics as well as in the domain of science. In addition, we found a smaller conceptual interference effect (in accuracy) in mathematicians in both tasks. The neural results showed that the left medial DLPFC was activated more strongly in incongruent than in congruent statements in mathematics, while in science we found activation in bilateral and medial DLPFC and the bilateral VLPFC. Contrary to our hypothesis, we did not find effects of mathematical expertise on brain activation in the mathematical task.

Conceptual interference in performance

Our finding of significant congruency effects in both domains and for both groups replicates the majority of all studies on conceptual knowl-edge representation, both in mathematics [6,27] and in science [3]. Our hypothesis regarding the effect of expertise was also corroborated. Mathematicians showed a smaller interference effect in accuracy than non-mathematicians, indicating that they are better at inhibiting naïve misconceptions in mathematics. Consistent with Shtulman and Har- rington [5] and Stricker et al. [28], individual differences in expertise were not related to the interference index in response time. This is in contrast to Allaire-Duquette et al. [8] and Lubin et al. [12], who found that the interference index in response time was also related to indi- vidual differences in expertise. However, in those two studies, the di- rection of the effect was different. While Lubin et al. found that the difference between incongruent and congruent items was smaller for experts than for novices, Allaire-Duquette et al. reported that the dif- ference in response time was larger for experts compared to novices. These contradicting findings may be due to methodological differences and/or due to differences in the domains and the level of expertise. While Lubin et al. used a negative priming task with mathematical items focusing on one specific misconception, Allaire-Duquette et al. applied the same speeded reasoning tasks we used, but in the domain of science. Further, whereas Lubin et al. had undergraduate students in mathematics as experts, Allaire-Duquette et al. used secondary school students with high competence as experts. Consequently, expertise was assessed with a measurement very closely related to the experimental task. In contrast, in our study, expertise was defined as having a large amount of university-level domain knowledge. and we focused on both mathematics and science from a broad perspective. Takeing all previous and especially the present findings together, it appears that the conceptual interference effect in accuracy in speeded reasoning tasks is more sensitive to individual knowledge differences than response times.

Even though we did not expect an effect of expertise in the domain of science, which we chose as a control domain, a lower interference effect in accuracy in the mathematicians, compared to the non-mathematicians, emerged. This suggests that mathematicians were better at inhibiting naïve misconceptions also in science. One possible explanation for this could be due to the sample as well as due to the item characteristics. It appears likely that individuals who decided to pursue a career in mathematics also have a higher competence in other related STEM fields, like physics or biology, which the science items draw on. However, this effect was smaller in the science (d 0.72) than in the mathematical task (d 1.24).

Thus, our behavioral findings demonstrate for the first time that mathematicians show a smaller conceptual interference effect over a broad range of mathematical sub-domains, suggesting that they can better deal with interfering naïve theories. However, only from the behavioral results, it is unclear if this is only due to their higher level of domain knowledge or whether they can better employ cognitive control mechanisms such as inhibiting the naïve concepts. The following neural results provide another level of analysis to answer this question.

= —

= —

Conceptual interference at the neural level: general effects

Based on fMRI conceptual interference studies in mathematics [1] and science [7,13] we expected the DLPFC, the VLPFC, and the ACC to be activated when inhibiting naïve theories. However, our hypothesis was only partly confirmed, and results differed depending on the domain.

For both mathematics and science, the correct evaluation of incongruent compared to congruent statements activated the same region in the left medial DLPFC (see Table A5; Fig. A2a). This result corresponds with the majority of studies on conceptual knowledge representation, which repeatedly found the DLPFC to be involved when inhibiting naïve concepts (e.g., [7,13]), and specifically with two recent studies [9,29] that found this particular part of the DLPFC (BA 8). According to Vaughn et al. [7], the proposed cognitive processes related to activation in the DLPFC are error detection, conflict monitoring, and inhibition. In mathematics, only the left medial DLPFC (BA 8) as well as the left precuneus (BA 7) was more strongly activated in the incongruent compared to the congruent condition. While in mathematics only these three activation clusters emerged in the incongruent > congruent contrast, in science the activation was more widespread.

In science, not only medial parts of the DLPFC, but also lateral parts of the DLPFC and many more activation clusters emerged. This broad activation in science is in line with Dumontheil et al. [29], who also found a widespread activation in the DLPFC. The distinction between medial and lateral parts of the DLPFC can also be seen in other studies. While Dumontheil et al. [29] found only lateral activation in the DLPFC (BA 8), Allaire-Duquette [9] also found medial activation of the DLPFC (BA 8). According to Taren et al. [30], the medial parts of the DLPFC are responsible for monitoring performance while the lateral parts are responsible for adjusting performance. Further, the medial DLPFC shows a posterior-to-anterior gradient that reflects complexity. While posterior parts are related to simple rules for control, more anterior parts are responsible for higher-order rules for control. From the three proposed functions (response conflict, decision conflict, strategy conflict), the medial DLPFC areas observed in the present study, both in mathematics and in science, are topographically closest to the areas responsible for decision conflict. In our case, participants had to decide between two conflicting theories – the naïve or the scientific one.

Further, in science, parts of the VLPFC (BA 47) were activated bilaterally, confirming our hypothesis. The proposed cognitive processes in both of these areas are executive functions as well as inhibition [7]. While several studies indicate that inhibitory control and visuospatial stimuli are lateralized right in the VLPFC, semantic inhibition and verbal stimuli activate the left VLPFC [31–33]. In addition, the left VLPFC shows functional distinctions between subregions. While the anterior part of the VLPFC (BA 47), which was activated in our case, is known to control access to stored conceptual representations, the mid VLPFC (BA 45) supports a domain-general-selection process after retrieval that resolves the competition between active representations [33]. The bilateral anterior part of the VLPFC was also activated in Dumontheil et al. [29]. These findings suggest that to correctly solve incongruent items in science, in general, more inhibition has to be applied than during the processing of congruent items, and especially cognitive control during the access to conceptual representations seems to be important.

While our hypothesis regarding the DLPFC and the VLPFC was mostly confirmed, contrary to our expectations, we found no significant activation for the incongruent > congruent contrast in the ACC, neither in mathematics nor in science. However, in contrast to both prefrontal regions, the ACC was not consistently reported in all relevant previous studies. Masson et al. (2014) found the ACC to be more strongly activated in experts than in novices in incongruent trials, but also in congruent trials. Brault Foisy et al. (2015) found the ACC only in novices but not in experts, Allaire-Duquette et al. (2021) only in the domain of biology but not in physics, and some studies did not find any involvement of the ACC at all [8,10,29]. Interestingly, some of the activations in BA 8, which we labeled as medial DLPFC seemed to be very close to, or just above the ACC. Nonetheless, this is not strong enough evidence to speak of an involvement of the ACC. Thus, the ACC does not seem to be as strongly tied to the resolution of conceptual interference as the DLPFC and the VLPFC.

The precuneus was the only additional region to be activated in the mathematics condition. Concentrating on the anatomical location of BA 7, this region is assumed to be involved during conflict monitoring, but also in reallocating attentional resources during visuospatial processing [7], episodic memory retrieval, self-processing, and consciousness [34]. Especially the function of conflict monitoring seems to be relevant, because in incongruent statements the naïve concepts are conflicting with the correct mathematical concepts, and this conflict has to be monitored and to be solved. The involvement of the precuneus to support conflict monitoring has also been observed in a study on complex causal thinking [35]. While this study used a conceptually different task, an incongruent condition, where theory and data were inconsistent, was compared to a congruent condition. Results showed a brain activation pattern associated with error detection and conflict monitoring, including the left DLPFC, ACC, and precuneus. Thus, the activation of the precuneus in the present study may indeed reflect the monitoring of a cognitive conflict and its resolution.

Additional significant brain regions in the science task were the left insula (BA 13), the angular gyrus bilaterally (BA 39), the left caudate (BA 48), the left anterior prefrontal cortex (BA 10), the left inferior temporal gyrus (BA 37), and the right cerebellum. The activation cluster in the left insula was located in the anterior region - a region that has also been reported by Potvin et al. [10] to be more strongly activated in the incongruent condition compared to the congruent condition in a speeded reasoning task in chemistry. In general, the insula has a high likelihood to be activated during cognitive processes and has been found to causally influence large scale brain networks like the central executive network and is assumed to be especially influential in tasks requiring greater cognitive control, as is the case when one has to inhibit naïve concepts. The central executive network additionally includes the DLPFC as well as the posterior parietal cortex and is engaged when performing a cognitively demanding task that requires attention [36]. As the angular gyrus (BA 39) is considered to be part of the posterior parietal cortex [37], this implies that the central executive network is activated while solving incongruent statements and overcoming naïve misconceptions in science. However, the angular gyrus is not only part of the central executive network, but also has been involved in conflict resolution [38]. While in less complex tasks mostly the right angular gyrus is activated, the left angular gyrus is only activated when a strong contextual/semantic conflict is present. As the conflict between naïve and scientific theories is a rather strong and semantic one, which was additionally enhanced by the item selection during the preliminary stages, the stronger activation in the incongruent condition both in the left and in the right angular gyrus is in line with previous findings. Further, the angular gyrus was also found by Dumontheil et al. [29] to be activated more strongly in the incongruent > congruent contrast.

The anterior prefrontal cortex (BA 10) on the left hemisphere was also significantly activated in the contrast incongruent > congruent. This region was also reported by several others to be involved in conceptual interference [1,8,14,16] and to be related to working memory as well as executive functions. Further, the left inferior temporal gyrus (BA 37) was activated, as was also the case in Dumontheil et al. [29]. While the inferior temporal gyrus is most well-known for visual object recognition, also decision making has been associated with this part of the brain [39]. Such decision-making processes may be involved in actively selecting answers in accordance with the scientific theory. Moreover, within the right cerebellum, a significant activation in Crus I and Crus II was observed, which is consistent with Allaire-Duquette et al. [9] and Dumontheil et al. [29].

In addition to obvious differences in the number and size of activation clusters for the incongruent > congruent contrast between the mathematics and science tasks, a direct contrast analysis revealed that the neural interference effect significantly differed between both domains. While the contrast incongruent > congruent was accompanied by stronger activation in the left DLPFC and VLPFC in science, in mathematics the inferior and superior parietal lobule bilaterally, including the precuneus, were more strongly activated (see Table A5; Fig. A2b). The stronger activation of the DLPFC and VLPFC may suggest that more inhibition is needed to overcome naïve concepts in science, which is unexpected as the behavioral interference effects in both domains are quite similar (η_p^2=.83 vs. 0.88). The stronger activation of the bilateral parietal areas for mathematics may be related to the material involved. Not only the precuneus but also the superior and inferior parietal lobules, are involved in mathematical problem solving [40–42]. This activation difference between mathematics and science does not correspond to the recent study from Dumontheil et al. [29] in which 34 adolescents worked on 48 statements each from the domains of science (biology, chemistry, and physics) and mathematics (number, algebra, ratio, geometry, probability, and statistics). They found no differences in brain activation between both domains for the incongruent > congruent contrast. However, it should be emphasized that in our study the brain regions were not completely dissimilar, in that in both domains participants activated the medial DLPFC (BA 8) to overcome naïve theories (see Fig. A1a and Table A3).

An answer to the question of why more inhibition is needed in science may be related to the type of involved concepts. Even though the statements from mathematics and science were constructed in the same way and were also matched for word count and sentence complexity, naïve theories in mathematics and science fundamentally differ. First, naïve concepts in science are learned much earlier via everyday experiences (e.g., a toddler learns about gravity by dropping a toy), while naïve concepts in mathematics are generally learned later, often with the beginning of formal education. Thus, one could argue that naïve concepts in science are more deeply anchored in the brain. Second, as has been pointed out by Potvin et al. [43] and also by Stricker et al. [6] the majority of statements in science refer to everyday perceptual experiences, whereas mathematical statements were more formal and mostly abstract. Third, the correctness of naïve mathematical theories seems to be context-dependent. While some naïve science concepts (e.g., the sun revolves around the sun) are simply scientifically incorrect, other naive mathematical concepts (e.g., natural number bias) may be incorrect only in some contexts (e.g., in fractions). While speculative at this point, all of these factors could be responsible for the observed differences in the neural effects.

Conceptual interference at the neural level: the impact of expertise

In addition to investigating the brain regions associated with conceptual interference in mathematics and science, another goal of the present study was to examine the moderating role of expertise. Results from other domains indicated that experts have a stronger activation in the DLPFC, VLPFC, and ACC than novices when solving incongruent items in their domain [14,16]. In addition, in one study on science, it was found that experts more strongly activated these regions in the incongruent > congruent contrast [8], indicating a greater recruitment of inhibitory control mechanisms. Contrary to our hypothesis, we did not find significant activation clusters when comparing both groups in the contrast incongruent > congruent in mathematics. We can only speculate why we did not find a moderating effect of mathematical expertise on conceptual interference at the neural level. To begin with, the effect in mathematics was not that pronounced, especially in comparison to science. Further, we used a more stringent correction for multiple comparisons in comparison to other studies [13,14,16] which found a moderating effect of expertise.

In the domain of science, both mathematicians and non-mathematicians significantly activated the left medial and lateral DLPFC more strongly during incongruent than during congruent statements (see Fig. A2a and Table A4). While a visual inspection indicated that mathematicians showed more distributed and larger activation clusters, especially in the left lateral DLPFC and left VLPFC, than non-mathematicians, supporting the assumption that they were better at inhibiting naïve theories, again a direct contrast analysis revealed no significant brain areas.

The visual inspection of the activation distribution in the contrast incongruent > congruent calculated separately for both groups in science (Table 4) is generally in line with Allaire-Duquette et al. [13], where experts activated the left DLPFC more strongly in incongruent than in congruent items, while novices showed no activation differences between incongruent and congruent. Further, the results correspond to the findings from Masson et al. [14] and Brault Foisy et al. [16], who found that experts showed significantly greater activation than novices in the left DLPFC when processing incongruent stimuli. While mathematicians showed significant activation in the left VLPFC (BA 47), the left caudate (BA 48), the left anterior prefrontal cortex (BA 10), the inferior temporal gyrus (BA 37) bilaterally, and the cerebellum bilaterally, non-mathematicians showed significant activation in the left insula (BA 13) and the left supplementary motor cortex (BA 6). As outlined above, in mathematicians these regions indicate that cognitive processes like executive functions, decision making, and attentional processing are involved when inhibiting naïve concepts. For non-mathematicians, the left insula is involved, suggesting that for them this is a task that needs great cognitive control. Further, in non-mathematicians, the left supplementary motor cortex (BA 6), more precisely the anterior part, showed significant activation. In previous studies, this region was more strongly activated for incongruent than for congruent statements by experts in their domain of expertise [10,16].

Limitations and future directions

One common problem in fMRI studies is the small number of participants resulting in low power. However, our sample (30 experts/31 novices) was larger than those in previous studies comparing experts and novices (12/12, Allaire-Duquette et al. [8]; 10/19, Brault Foisy et al. [16]; 11/12, Masson et al. [14]), and also larger than those only focusing on experts (25, Allaire-Duquette et al. [9]; 17, Potvin et al. [10]). While for the most part low statistical power does not seem to be a major limitation in the present study, there are a few other limitations that need to be mentioned. First, even though we defined our mathematicians as experts, they were mostly still bachelor students. Nonetheless, they had at least one and a half years of university-level math education and differed substantially in their mathematical knowledge from novices. However, while we did see substantial behavioral differences of expertise in the mathematical task, we did not find such differences at the neural level. Choosing experts with a higher level of mathematical expertise (i.e., holding a PhD) for a further study may provide more differentiated information. Second, another common problem in fMRI studies, which was also the case in our study, is that the interpretation of fMRI results involves reverse inference [44]. In reverse inference, one concludes from a specific pattern of brain activation that a particular cognitive process (e.g., inhibition) is recruited by an experimental task. This has to be kept in mind when reading the interpretation of our neuronal results (for more information on reverse inference in conceptual knowledge representation see Allaire-Duquette et al. [9]). Third, even though the mathematical statements were developed by experts in mathematics education, there might be some ambiguities. One could argue whether the answering tendency for some statements (e.g., 20 mm^2 are larger than 1 cm^2) is really driven by an underlying naïve theory (e.g., larger numbers always mean that something is larger) or if the interference effect is only due to more salient information (one focuses only on the numbers instead of on the metrics because numbers are more protruding than metrics).

Conclusion

This study aimed to bridge the gaps between previous research on conceptual knowledge representation in mathematics, the effects of expertise, and related brain activation. Our study revealed that naïve misconceptions in mathematics are still present even in experts, and formal instruction cannot entirely eradicate those misconceptions.

However, experts in mathematics showed a smaller interference effect in accuracy, suggesting that they can better inhibit the naïve theories. The fMRI results corroborate these findings by demonstrating that brain regions associated with inhibitory control are involved in overcoming naive theories. In addition, they suggest that the extent of inhibitory control processes differs between the domains of mathematics and science, which may be due to fundamental differences in the involved concepts. However, an impact of mathematical expertise on the brain activation in conceptual interference was not visible, neither in mathematics nor in science. Overall, the present study provided further insights into the neural and cognitive processes underlying interference effects in conceptual knowledge.

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

6. Appendix

All statements in the domain of mathematics

All statements in the domain of science

Additional fMRI analyses

To statistically test similarities and differences in interference related brain activations between groups we calculated conjunction and contrast analysis using the Matlab Toolbox GLM_Flex_Fast4 [24]. In the conjunction analyses we tested which areas are activated in both groups. To this end, we used the logical AND function of the NumPy Python package to generate a binary mask between two contrasts. This mask was then used to extract the cluster sizes, coordinates and peak values from the T-maps of these contrasts. In the contrast analysis we calculated a new first level analyses for each participant with FitLins, directly contrasting incongruent > congruent in each domain (instead of first contrasting each condition with a baseline, as has been done in the other analyses). We then estimated a second-level random effects whole-brain model comparing both groups. All reported t-statistics were family-wise error (FWE) corrected with a threshold at $p_{FWE} < 0.001$ at voxel level. The threshold correction and the figures were done using the Nilearn toolbox in Python. For the automatic labeling of the brain regions and the Brodmann Areas we used the label4MRI R package [25] which uses the AAL (automatic anatomical labeling) atlas [26].

To statistically test similarities and differences in interference related brain activations between domains, the same conjunction and contrast analyses as describes above, only now for domains and not for groups, were calculated.

Similarities and differences in interference related brain activations between groups

In mathematics no significant activation similarities or differences in the contrast incongruent > congruent between the groups were found (see Table A3).

In science no significant activation differences in the contrast incongruent > congruent between the groups were found (see Table A4). An exploratory conjunction analysis revealed that both mathematicians and non-mathematicians demonstrated significant activation in the lateral and medial DLPFC (see Fig. A1 and Table A4 in the appendix).

Table A1

All statements measuring interference of naïve and mathematical theories.

Fractions **Subtopic**	**Statement**
Addition von Brüchen	1/10 + 1/10 = 1/5
	1/3 + 1/4 = 2/7
	1/4 + 1/4 = 1/4
	1/4 + 2/4 = 3/4
Bedeutung von Bruch als Anteil oder Verhältnis	3 von 4 kann bedeuten: 3 von 4 Perlen sind schwarz.
	3 von 4 kann bedeuten: 3 von 7 Perlen sind schwarz.
	3 zu 4 kann bedeuten: 3 von 4 Perlen sind schwarz.
	3 zu 4 kann bedeuten: 3 von 7 Perlen sind schwarz.
Dichte von Bruchzahlen	Es gibt Zahlen zwischen 1/10 und 1/10.
	Es gibt Zahlen zwischen 1/10 und 10/100.
	Es gibt Zahlen zwischen 1/5 und 2/5.
	Es gibt Zahlen zwischen 1/5 und 3/5.
Division verkleinert immer	4 : 1/4 ist größer als 4.
	4 : 1/4 ist kleiner als 4.
	4 : 4/1 ist größer als 4.
	4 : 4/1 ist kleiner als 4.
Größenvergleich	1/4 ist größer als 1/2.
	1/5 ist größer als 1/6.
	1/9 ist größer als 1/8.
	3/4 ist größer als 1/2.
Multiplikation mit Erweitern verwechseln	6 ist das Dreifache von 2.
	6/9 ist das Dreifache von 2/3.
	6/9 ist das Dreifache von 2/9.
	9 ist das Dreifache von 2.
Multiplikation von Brüchen (Von-Deutung)	Die Hälfte von 2/3 ist dasselbe wie 1/2 + 2/3.
	Die Hälfte von 2/3 ist dasselbe wie 1/2 * 2/3.
	Die Hälfte von 2/3 ist dasselbe wie 2/3 : 1/2.
	Die Hälfte von 2/3 ist dasselbe wie 2/3 : 2.
Umwandlung zwischen verschiedenen Zahlbereichen und Einbettung der Zahlbereiche	Alle Brüche mit Bruchstrich können auch als Dezimalbrüche ausgedrückt werden.
	Alle ganzen Zahlen können auch als Brüche ausgedrückt werden.
	Alle Dezimalbrüche können auch als Brüche mit Bruchstrich ausgedrückt werden.
	Alle Brüche können auch als ganze Zahlen ausgedrückt werden.
Algebra	
Subtopic	**Statement**
Distributivgesetze übergeneralisiert	$a + (b + c) = a * (b * c)$
	$a + (b * c) = (a + b) * (a + c)$
	$a * (b + c) = a * b + a * c$
	$a * (b + c) = a * c + a * b$
Null als Zahl	0 ist eine Zahl
	0 ist keine Zahl
	1 ist eine Zahl
	1 ist keine Zahl
Potenzen mit konkreten Zahlen	$2^2 = 4$
	$2^3 = 5$
	$2^3 = 6$
	$2^3 = 8$
Potenzieren mit 0	$0^2 = 0$
	$2^{\circ} = 0$
	$2^{\circ} = 1$
	$2^{\circ} = 4$
Ungültigkeit des Assoziativgesetzes bei der Subtraktion	20 - (8 - 7) = (20 - 8) + 7
	20 - (8 - 7) = (20 - 8) - 7
	20 - (8 - 7) = 20 - (9 - 7)
	20 - (8 - 7) = 20 - (9 - 8)
Units and Geometry	
Subtopics	**Statement**
Größere (ähnliche) Form bedeutet auch größere Winkel	Wird ein Dreieck vergrößert, bleibt die Anzahl der Winkel gleich.

(*continued on next page*)

Table A1 (*continued*)

Fractions **Subtopic**	**Statement**
	Wird ein Dreieck vergrößert, bleibt die Größe der Winkel gleich. Wird ein Dreieck vergrößert, wird die Anzahl der Winkel verändert. Wird ein Dreieck vergrößert, wird die Größe der Winkel verändert.
Keine Begriffsinkulsion von Quadrat als Rechteck	Jedes Quadrat ist ein Rechteck. Jedes Quadrat ist ein Viereck. Nicht jedes Quadrat ist ein Rechteck. Nicht jedes Quadrat ist ein Viereck.
Quadrat=Viereck	Alle Vierecke haben rechte Winkel. Alle Vierecke haben vier Winkel. Einige Vierecke haben sieben Winkel. Einige Vierecke haben spitze Winkel.
Umrechnungsschwierigkeiten bzgl. der Flächeneinheiten	20 mm^2 sind kleiner als 10 cm^2. 20 mm^2 sind größer als 1 cm^2. 20 mm^2 sind größer als 10 cm^2. 20 mm^2 sind kleiner als 1 cm^2.
Umrechnungsschwierigkeiten bzgl. der Volumina	200 cm^3 sind mehr als 1 m^3. 200 cm^3 sind mehr als 100 m^3. 200 cm^3 sind weniger als 1 m^3. 200 cm^3 sind weniger als 100 m^3.
Umrechnungsschwierigkeiten bzgl. der Zeiteinheiten	1,59 h sind mehr als 1 h und weniger als 1½ h. 1,59 h sind mehr als 1½ h und weniger als 2 h. 1,59 h sind mehr als 1¾ h und weniger als 2 h. 1,59 h sind mehr als 1 h und weniger als 1¾ h.
Verwechslung von Form und Körper	Eine 1€ Münze ist ein Geldstück. Eine 1€ Münze ist ein Kreis. Eine 1€ Münze ist ein Quader. Eine 1€ Münze ist ein Zylinder.
Verwechslung von Quader und Quadrat	Ein Quader hat genau 4 Ecken. Ein Quader hat genau 8 Ecken. Ein Quadrat hat genau 4 Ecken. Ein Quadrat hat genau 8 Ecken.
Basic concepts	
Subtopic	**Statement**
Lautgetreue Schreibweise beim Transkribieren von Zahlen	Die Zahl dreitausendsechsundsiebzig hat mehr als vier Nullen. Die Zahl dreitausendsechsundsiebzig hat mehr als zwei Nullen. Die Zahl dreitausendsechsundsiebzig hat weniger als vier Nullen. Die Zahl dreitausendsechsundsiebzig hat weniger als zwei Nullen.
Logarithmischer vs. linearer Zahlenstrahl	Die Zahl 275 Mio. liegt auf einem Zahlenstrahl von 0 bis 1 Mrd. außerhalb des Zahlenstrahls. Die Zahl 275 Mio. liegt auf einem Zahlenstrahl von 0 bis 1 Mrd. innerhalb des Zahlenstrahls. Die Zahl 275 Mio. liegt auf einem Zahlenstrahl von 0 bis 1 Mrd. näher an 0 als an 1 Mrd. Die Zahl 275 Mio. liegt auf einem Zahlenstrahl von 0 bis 1 Mrd. näher an 1 Mrd. als an 0.
Wechsel der Rechenrichtung, wenn es ins Auge springt: Division	Das Ergebnis von 300 : 60 ist genau 20. Das Ergebnis von 300 : 60 ist genau 60. Das Ergebnis von 300 : 60 ist kleiner als 20. Das Ergebnis von 300 : 60 ist kleiner als 60.
Wechsel der Rechenrichtung, wenn es ins Auge springt: Subtraktion	Das Ergebnis von 64 - 25 ist genau 41. Das Ergebnis von 64 - 25 ist genau 51. Das Ergebnis von 64 - 25 ist kleiner als 41. Das Ergebnis von 64 - 25 ist kleiner als 51.

Table A2

All statements measuring interference of naïve and scientific theories.

Astronomy **Subtopic**	**Statement**
Stern	Die Sonne erzeugt Licht Die Sonne erzeugt Klang. Der Mond erzeugt Licht. Die Sonne erzeugt Anziehungskraft.
Mondphase	Die Mondphasen werden durch Veränderung der Lichtintensität verursacht. Die Mondphasen werden durch Wolken verursacht. Die Mondphasen werden durch den Schatten der Erde verursacht. Die Mondphasen werden durch die Umlaufbahn des Mondes verursacht.
Evolution	
Subtopic	**Statement**
Gemeinsame Abstammung	Menschen sind näher mit Menschenaffen verwandt als mit anderen Affen. Wale sind näher mit Pflanzen verwandt als Fische. Menschenaffen sind näher mit anderen Affen verwandt als Menschen. Wale sind näher als Fische mit Menschen verwandt.
Variation	Evolution erfordert das Überleben der am besten angepassten Individuen. Evolution erfordert ein stabiles Klima. Evolution erfordert lange Zeitperioden. Evolution erfordert Variation innerhalb von Tierarten.
Selektion	Die meisten Organismen sind an ihre Umgebung angepasst. Die meisten Organismen leben in einem *gem*äßigten Klima. Die meisten Organismen haben viel zu essen. Die meisten Organismen sterben bevor sie Nachkommen hinterlassen.
Adaption	Biologische Spezies unterliegen der Evolution. Leblose Objekte unterliegen der Evolution. Individuelle Organismen unterliegen der Evolution. Computerviren unterliegen der Evolution.
Genetics	
Subtopic	**Statement**
Erblichkeit	Die Haarfarbe ist erblich. Gepiercte Ohren sind erblich. Immunität gegen Windpocken ist erblich. Intelligenz ist erblich.
Germs	
Subtopic	**Statement**
Kontamination	Verdorbenes Fleisch enthält Keime. Sonnenlicht enthält Keime. Urin enthält Keime. Spülschwämme enthalten Keime.
Infektion	Keime können durch Schnitte in den Körper gelangen. Keime können durch Haare in den Körper gelangen. Keime können durch die Haut in den Körper gelangen. Keime können durch die Augen in den Körper gelangen.
Mikroben	Ein Keim hat eine Form. Ein Keim hat Gefühle. Ein Keim hat einen Geruch. Ein Keim hat eine DNA.
Matter	
Subtopic	**Statement**
Dichte	Stahl ist dichter als Schaumstoff. Schaumstoff ist dichter als ein Ziegelstein. Eis ist dichter als Wasser. Eine kalte Münze ist dichter als eine heiße Münze.
Mechanics	
Subtopic	**Statement**
Schwerkraft	Ein Amboss fällt schneller durch die Luft als eine Feder. Helle Objekte fallen schneller durch die Luft als dunkle. Schwere Kugeln fallen schneller durch die Luft als leichte. Spitze Objekte fallen schneller durch die Luft als abgeflachte.
Mechanics	
Subtopic	**Statement**
Leben	Fische sind lebendig. Steine sind lebendig. Die Sonne ist lebendig. Korallen sind lebendig.

(*continued on next page*)

Table A2 (*continued*)

Astronomy **Subtopic**	**Statement**
Tod	Schildkröten können sterben.
	Schraubenzieher können sterben.
	Wolken können sterben.
	Pilze können sterben.
Reproduktion	Tiger können sich fortpflanzen.
	Stühle können sich fortpflanzen.
	Raupen können sich fortpflanzen.
	Farne können sich fortpflanzen.
Metabolismus	Menschen wandeln Nahrung in Energie um.
	Steine wandeln Nahrung in Energie um.
	Pflanzen wandeln Nahrung in Energie um.
	Bakterien wandeln Nahrung in Energie um.
Verwandtschaft	Ein Baby kann eine Nichte oder ein Neffe sein.
	Ein Baby kann eine Mutter oder ein Vater sein.
	Ein Baby kann sein eigenes älteres Geschwister sein.
	Ein Baby kann eine Tante oder ein Onkel sein.
Thermodynamics	
Subtopic	**Statement**
Wärme	Die Sonne hat Wärme.
	Schwerkraft hat Wärme.
	Ein Atom hat Wärme.
	Eis hat Wärme.
Wärmequelle	Öfen produzieren Wärme.
	Regen produziert Wärme.
	Mäntel produzieren Wärme.
	Druck produziert Wärme.
Wärmetransfer	Wärme überträgt sich von Wasser auf Eis.
	Wärme überträgt sich von kalten Objekten auf warme Objekte.
	Wärme überträgt sich von Wasser auf Dampf.
	Wärme überträgt sich von kalten Objekten auf noch kältere Objekte.
Temperatur	Eis ist kälter als Wasser.
	Dampf ist kälter als Eis.
	Zwei Becher Eis sind kälter als einer.
	Kochendes Wasser ist kälter als sein Dampf.
Wärmeausdehnung	Wärme erhöht die Temperatur eines Objektes.
	Wärme verstärkt die Farbe eines Objektes.
	Wärme erhöht das Gewicht eines Objektes.
	Wärme erhöht die Größe eines Objektes.
Waves	
Subtopic	**Statement**
Farbe	Rote Objekte reflektieren rotes Licht.
	Rote Objekte reflektieren blaues Licht.
	Rote Objekte absorbieren rotes Licht.
	Rote Objekte absorbieren blaues Licht.
Klang	Töne können laut oder leise sein.
	Töne können tot oder lebendig sein.
	Töne können nah oder fern sein.
	Töne können direkt oder reflektiert sein.
Reflektion	Spiegel reflektieren Licht.
	Schaumstoff reflektiert Töne.
	Prismen reflektieren Licht.
	Flächen reflektieren Töne.

Table A3

Conjunction (A) and Contrast Analysis (B) for the contrast incongruent > congruent in mathematics for mathematicians and non-mathematicians.

Brain Region	AAL Label	Brodmann Areal (BA)	MNI X	Y	Z	Cluster size *k*
Mathematics: Mathematicians Incongruent > Congruent AND Non-mathematicians Incongruent > Congruent						
None						
Mathematics: Mathematicians Incongruent > Congruent ≠ Science: Non-mathematicians Incongruent > Congruent						
Mathematicians > Non mathematicians						
None						
Non mathematicians > Mathematicians						
None						

All comparisons are reported at FWE corrected threshold at $p_{FWE} < 0.001$ at voxel level.

Table A4

Conjunction (A) and Contrast Analysis (B) for the contrast incongruent > congruent in science for mathematicians and non-mathematicians.

Brain Region	AAL Label	Brodmann Areal (BA)	MNI X	Y	Z	Cluster size *k*
(A) Science: Mathematicians Incongruent > Congruent AND Non-mathematicians Incongruent > Congruent						
Lateral DLPFC	Frontal_Mid_L	Left-BA8	—27	20	52	384
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	—7	28	40	184
(B) Science: Mathematicians Incongruent > Congruent ≠ Science: Non-mathematicians Incongruent > Congruent						
Mathematicians > Non mathematicians						
None						
Non mathematicians > Mathematicians						
None						

All comparisons are reported at FWE corrected threshold at $p_{FWE} < 0.001$ at voxel level.

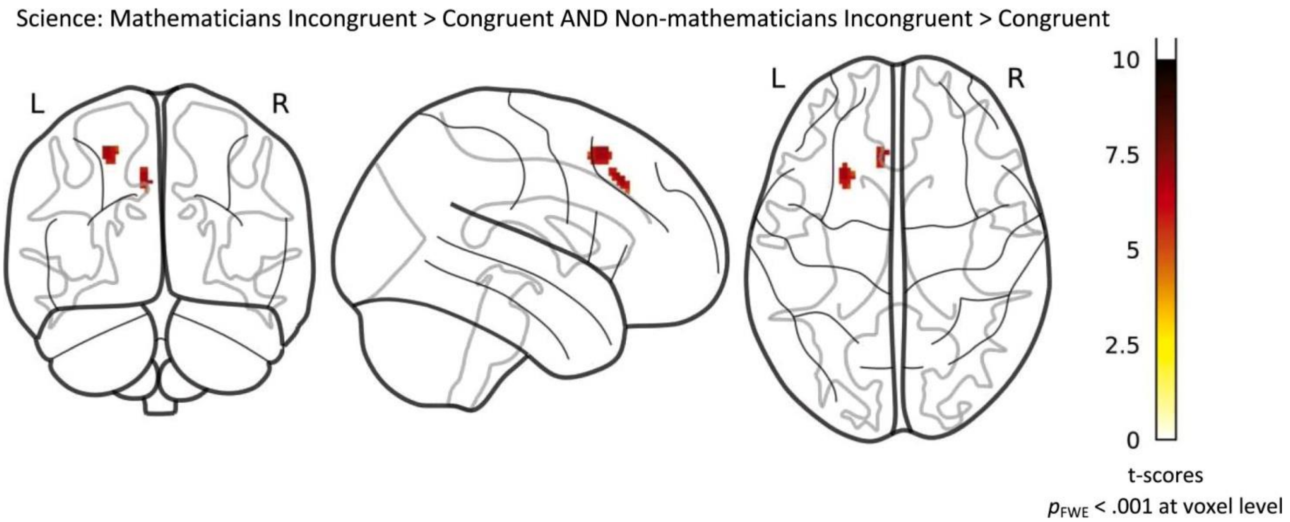


Fig. A1. Science: Mathematicians Incongruent > Congruent AND Non-mathematicians Incongruent > Congruent.

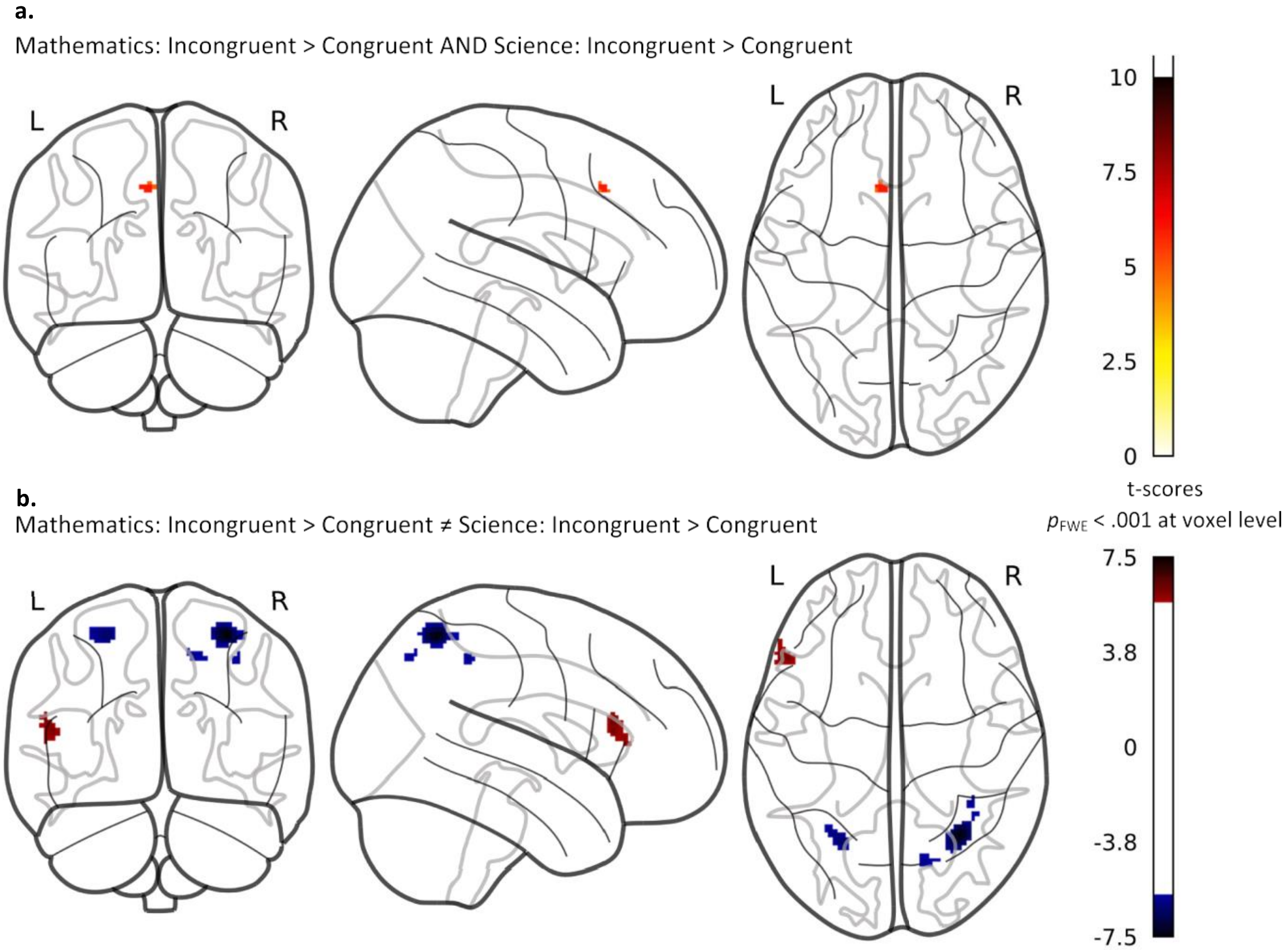


Fig. A2a. Mathematics: Incongruent > Congruent AND Science: Incongruent > Congruent. B. Mathematics: Incongruent > Congruent ≠ Science: Incongruent > Congruent.

Table A5

Conjunction (A) and Contrast Analysis (B) for the contrast incongruent > congruent for both domains.

Brain Region	AAL Label	Brodmann Areal (BA)	MNI X	Y	Z	Cluster size *k*
(A) **Mathematics: Incongruent > Congruent AND Science: Incongruent > Congruent**						
Medial DLPFC	Frontal_Sup_Medial_L	Left-BA8	—7	22	44	96
(A) **Mathematics: Incongruent > Congruent ≠ Science: Incongruent > Congruent**						
Mathematics > Science (blue)						
Superior parietal lobule	Parietal_Sup_R	Right-BA7	32	—59	60	1024
Superior parietal lobule	Parietal_Sup_L	Left-BA7	—27	—61	60	496
Inferior parietal lobule	Parietal_Inf_R	Right-BA7	38	—43	48	72
Precuneus	Precuneus_R	Right-BA7	16	—71	50	128
Superior parietal lobule	Parietal_Sup_R	Right-BA7	38	—51	60	40
Science > Mathematics (red)						
DLPFC	Frontal_Inf_Tri_L	Left-BA46	—55	26	20	304
VLPFC	Frontal_Inf_Tri_L	Left-BA45	—55	32	10	
VLPFC	Frontal_Inf_Tri_L	Left-BA45	—53	30	10	8

All comparisons are reported at FWE corrected threshold at $p_{\text{FWE}} < 0.001$ at voxel level.

Similarities and differences in interference related brain activations between domains

Both domains activated the medial DLPFC (Fig. A2a and Table A5), while science statements activated the left DLPFC and VLPFC stronger, and mathematical statements the inferior and superior parietal lobule bilaterally (Fig. A2b and Table A5).

References

[1] R. Stavy, R. Babai, Overcoming intuitive interference in mathematics: insights from behavioral, brain imaging and intervention studies, ZDM 42 (2010) 621–633, https://doi.org/10.1007/s11858-010-0251-z.

[2] X. Vamvakoussi, W. Van Dooren, L. Verschaffel, Brief Report. Educated adults are still affected by intuitions about the effect of arithmetical operations: evidence from a reaction-time study, Educ. Stud. Math. 82 (2013) 323–330, https://doi.org/10.1007/s10649-012-9432-8.

[3] L. Mason, S. Zaccoletti, Inhibition and Conceptual Learning in Science: a Review of Studies, Educ. Psychol. Rev. 33 (2021) 181–212, https://doi.org/10.1007/s10648-020-09529-x.

[4] A. Shtulman, J. Valcarcel, Scientific knowledge suppresses but does not supplant earlier intuitions, Cognition 124 (2012) 209–215, https://doi.org/10.1016/j.cognition.2012.04.005.

[5] A. Shtulman, K. Harrington, Tensions Between Science and Intuition Across the Lifespan, Top. Cogn. Sci. 8 (2016) 118–137, https://doi.org/10.1111/tops.12174.

[6] J. Stricker, S.E. Vogel, S. Schöneburg-Lehnert, T. Krohn, S. Dögnitz, N. Jud, M. Spirk, M.-C. Windhaber, M. Schneider, R.H. Grabner, Interference between naïve and scientific theories occurs in mathematics and is related to mathematical achievement, Cognition 214 (2021), 104789, https://doi.org/10.1016/j.cognition.2021.104789.

[7] A.R. Vaughn, R.D. Brown, M.L. Johnson, Understanding conceptual change and science learning through educational neuroscience, Mind, Brain, Educ 14 (2020) 82–93, https://doi.org/10.1111/mbe.12237.

[8] G. Allaire-Duquette, M. Bélanger, R.H. Grabner, K. Koschutnig, S. Masson, Individual differences in science competence among students are associated with ventrolateral prefrontal cortex activation, J. Neurosci. Res. 97 (2019) 1163–1178, https://doi.org/10.1002/jnr.24435.

[9] G. Allaire-Duquette, L.-M. Brault Foisy, P. Potvin, M. Riopel, M. Larose, S. Masson, An fMRI study of scientists with a Ph.D. in physics confronted with naive ideas in science, Npj Sci. Learn. 6 (2021) 11, https://doi.org/10.1038/s41539-021-00091-x.

[10] P. Potvin, G. Malenfant-Robichaud, C. Cormier, S. Masson, Coexistence of misconceptions and scientific conceptions in chemistry professors: a mental chronometry and fMRI study, Front. Educ. 5 (2020) 1–16, https://doi.org/10.3389/feduc.2020.542458.

[11] A. Obersteiner, W. Van Dooren, J. Van Hoof, L. Verschaffel, The natural number bias and magnitude representation in fraction comparison by expert mathematicians, Learn. Instr. 28 (2013) 64–72, https://doi.org/10.1016/j.learninstruc.2013.05.003.

[12] A. Lubin, S. Rossi, C. Lanoë, J. Vidal, O. Houdé, G. Borst, Expertise, inhibitory control and arithmetic word problems: a negative priming study in mathematics experts, Learn. Instr. 45 (2016) 40–48, https://doi.org/10.1016/j.learninstruc.2016.06.004.

[13] L. Nenciovici, G. Allaire-Duquette, S. Masson, Brain activations associated with scientific reasoning: a literature review, Cogn. Process. 20 (2019) 139–161, https://doi.org/10.1007/s10339-018-0896-z.

[14] S. Masson, P. Potvin, M. Riopel, L.-M.B. Foisy, Differences in brain activation between novices and experts in science during a task involving a common misconception in electricity, Mind, Brain, Educ. 8 (2014) 44–55, https://doi.org/10.1111/mbe.12043.

[15] P. Potvin, É. Turmel, S. Masson, Linking neuroscientific research on decision making to the educational context of novice students assigned to a multiple-choice scientific task involving common misconceptions about electrical circuits, Front. Hum. Neurosci. 8 (2014) 14, https://doi.org/10.3389/fnhum.2014.00014.

[16] L.-M. Brault Foisy, P. Potvin, M. Riopel, S. Masson, Is inhibition involved in overcoming a common physics misconception in mechanics? Trends Neurosci. Educ. 4 (2015) 26–36, https://doi.org/10.1016/j.tine.2015.03.001.

[17] A.O. Jäger, H.-.M. Süß, A. Beauducel, Berliner Intelligenzstruktur-Test [Berlin intelligence structure test], (1997).

[18] F. Jasper, D. Wagener, Mathematiktest Für Die Personalauswahl : M-PA, Hogrefe, Göttingen, 2013. https://madoc.bib.uni-mannheim.de/34304/.

[19] Authors et al. (2021).

[20] J. Peirce, R. Hirst, M. MacAskill, Building EXperiments in PsychoPy, 2nd Editio, SAGE Publications Ltd, 2022.

[21] IBM Corp, IBM SPSS Statistics for Windows, (2021).

[22] O. Esteban, C.J. Markiewicz, R.W. Blair, C.A. Moodie, A.I. Isik, A. Erramuzpe, J. D. Kent, M. Goncalves, E. DuPre, M. Snyder, H. Oya, S.S. Ghosh, J. Wright, J. Durnez, R.A. Poldrack, K.J. Gorgolewski, fMRIPrep: a robust preprocessing pipeline for functional MRI, Nat. Methods 16 (2019) 111–116, https://doi.org/10.1038/s41592-018-0235-4.

[23] C.J. Markiewicz, A. De La Vega, A. Wagner, Y.O. Halchenko, K. Finc, R. Ciric, M. Goncalves, D.M. Nielson, J.D. Kent, J.A. Lee, S. Bansal, R.A. Poldrack, K.J. Gorgolewski, poldracklab/fitlins: 0.10.1, (2022). 10.5281/ZENODO.6397685.

[24] GLM_Flex_Fast4, (n.d.). https://habs.mgh.harvard.edu/researchers/data-tools/glm-flex-fast2/.

[25] label4MRI, (n.d.). https://github.com/yunshiuan/label4MRI.

[26] E.T. Rolls, C.-C. Huang, C.-P. Lin, J. Feng, M. Joliot, Automated anatomical labelling atlas 3, Neuroimage 206 (2020), 116189, https://doi.org/10.1016/j.neuroimage.2019.116189.

[27] A. Brookman-Byrne, D. Mareschal, A.K. Tolmie, I. Dumontheil, Inhibitory control and counterintuitive science and maths reasoning in adolescence, PLoS ONE 13 (2018) 1–20, https://doi.org/10.1371/journal.pone.0198973.

[28] M. Schneider, S. Merz, J. Stricker, B. De Smedt, J. Torbeyns, L. Verschaffel, K. Luwel, Associations of number line estimation with mathematical competence: a meta-analysis, Child Dev. 00 (2018) 1–18, https://doi.org/10.1111/cdev.13068.

[29] I. Dumontheil, A. Brookman-Byrne, A.K. Tolmie, D. Mareschal, Neural and cognitive underpinnings of counterintuitive science and math reasoning in adolescence, J. Cogn. Neurosci. (2022) 1–25, https://doi.org/10.1162/jocn_a_01854.

[30] A.A. Taren, V. Venkatraman, S.A. Huettel, A parallel functional topography between medial and lateral prefrontal cortex: evidence and implications for cognitive control, J. Neurosci. 31 (2011) 5026–5031, https://doi.org/10.1523/JNEUROSCI.5762-10.2011.

[31] M.A.L. Ralph, E. Jefferies, K. Patterson, T.T. Rogers, The neural and computational bases of semantic cognition, Nat. Rev. Neurosci. 18 (2017) 42–55, https://doi.org/10.1038/nrn.2016.150.

[32] B.J. Levy, A.D. Wagner, Cognitive control and right ventrolateral prefrontal cortex: reflexive reorienting, motor inhibition, and action updating, Ann. NY Acad. Sci. 1224 (2011) 40–62, https://doi.org/10.1111/j.1749-6632.2011.05958.x.

[33] D. Badre, A.D. Wagner, Left ventrolateral prefrontal cortex and the cognitive control of memory, Neuropsychologia 45 (2007) 2883–2901, https://doi.org/10.1016/j.neuropsychologia.2007.06.015.

[34] A.E. Cavanna, M.R. Trimble, The precuneus: a review of its functional anatomy and behavioural correlates, Brain 129 (2006) 564–583, https://doi.org/10.1093/brain/awl004.

[35] J.A. Fugelsang, K.N. Dunbar, Brain-based mechanisms underlying complex causal thinking, Neuropsychologia 43 (2005) 1204–1213, https://doi.org/10.1016/j.neuropsychologia.2004.10.012.

[36] M.D. Fox, M. Corbetta, A.Z. Snyder, J.L. Vincent, M.E. Raichle, Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems, Proc. Natl. Acad. Sci. 103 (2006) 10046–10051, https://doi.org/10.1073/pnas.0604187103.

[37] J.R. Whitlock, Posterior parietal cortex, Curr. Biol. 27 (2017) R691–R695, https://doi.org/10.1016/j.cub.2017.06.007.

[38] M.L. Seghier, The Angular Gyrus, Neurosci 19 (2013) 43–61, https://doi.org/10.1177/1073858412440596.

[39] Y.-.H. Lin, I.M. Young, A.K. Conner, C.A. Glenn, A.R. Chakraborty, C.E. Nix, M. Y. Bai, V. Dhanaraj, R.D. Fonseka, J. Hormovas, O. Tanglay, R.G. Briggs, M. E. Sughrue, Anatomy and white matter connections of the inferior temporal gyrus, World Neurosurg. 143 (2020) e656–e666, https://doi.org/10.1016/j.wneu.2020.08.058.

[40] M. Amalric, S. Dehaene, A distinct cortical network for mathematical knowledge in the human brain, Neuroimage 189 (2019) 19–31, https://doi.org/10.1016/j.neuroimage.2019.01.001.

[41] M. Arsalidou, M.J. Taylor, Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations, Neuroimage 54 (2011) 2382–2393, https://doi.org/10.1016/j.neuroimage.2010.10.009.

[42] Z. Hawes, H.M. Sokolowski, C.B. Ononye, D. Ansari, Neural underpinnings of numerical and spatial cognition: an fMRI meta-analysis of brain regions associated with symbolic number, arithmetic, and mental rotation, Neurosci. Biobehav. Rev. 103 (2019) 316–336, https://doi.org/10.1016/j.neubiorev.2019.05.007.

[43] P. Potvin, S. Masson, S. Lafortune, G. Cyr, Persistence of the intuitive conception that heavier objects sink more: a reaction time study with different levels of interference, Int. J. Sci. Math. Educ. 13 (2015) 21–43, https://doi.org/10.1007/s10763-014-9520-6.

[44] R.A. Poldrack, Can cognitive processes be inferred from neuroimaging data? Trends Cogn. Sci. 10 (2006) 59–63, https://doi.org/10.1016/j.tics.2005.12.004.